Signal Processing 29 (1992) 113–140
Elsevier

A unifying view of coherence in signal processing*

William A. Gardner (Member EURASIP)
Department of Electrical Engineering and Computer Science, University of California, Davis, CA 95616, USA

Received 22 August 1990
Revised 13 August 1991

Abstract. The concept of *coherence* is fundamental and quite important in all fields dealing with fluctuating quantities. Although there is a commonality among the many uses of the term coherence, the precise meaning of this term seems to vary from one field to another and even within some fields. The purpose of this tutorial paper is to present a unifying view of the concept of coherence that is particularly relevant to the related fields of statistical signal processing and time-series analysis, which permeate numerous other more specialized fields. Precise mathematical definitions of a variety of types of coherence are given and are related to commonly used physical meanings. A brief survey of methods, which can be implemented with digital signal processing algorithms, for exploiting the various types of coherence that can occur in measurements of fluctuating quantities is presented. Of the three basic types of coherence – temporal, spectral and spatial – some emphasis is placed on spectral coherence since its use in signal processing has received the least attention in the literature.

Zusammenfassung. Das Konzept der *Kohärenz* ist fundamental und sehr wichtig auf allen Gebieten, bei denen es um veränderliche Größen geht. Obwohl es eine Gemeinsamkeit in den vielfältigen Anwendungen des Begriffs Kohärenz gibt, variiert die genaue Bedeutung dieses Ausdrucks in den unterschiedlichen Anwendungsbereichen und sogar innerhalb einzelner Bereiche. Das Ziel der vorliegenden Übersichtsarbeit besteht darin, eine einheitliche Darstellung des Konzeptes der Kohärenz in Hinblick auf statistische Signalverarbeitung und Zeitfolgen-Analyse zu geben, womit auch zahlreiche andere speziellere Gebiete erfaßt werden. Präzise mathematische Definitionen einer Vielfalt von Typen von Kohärenz werden gegeben und mit den gemeinhin benutzten physikalischen Bedeutungen verglichen. Es wird ein kurzer Überblick über die Anwendungsmöglichkeiten der verschiedenen Typen von Kohärenz wiedergegeben, die sich im Bereich der Messung veränderlicher Größen anbieten. Von den drei grundsätzlichen Typen von Kohärenz – zeitlich, spektral und räumlich – wird das Gewicht auf die spektrale Kohärenz gelegt, da ihre Anwendung bisher die geringste Aufmerksamkeit in der Literatur erfahren hat.

Résumé. Le concept de *cohérence* est fondamental et tout à fait important dans tous les domaines traitant de quantités fluctuantes. Bien qu'il y ait un principe commun aux usages nombreux du terme cohérence, la signification précise de ce terme semble varier quelque peu d'un domaine à un autre et même à l'intérieur de certains domaines. Le but de cet article à caractère pédagogique est de présenter une vue unificatrice du concept de cohérence particulièrement adaptée aux domaines apparentés du traitement statistique du signal et de l'analyse de séries temporelles, qui ont une influence sur bon nombre d'autres domaines plus spécialisés. Des définitions mathématiques précises d'une variété de types de cohérence sont données et reliées aux interprétations physiques couramment utilisées. Un bref rappel des façons possibles d'utiliser les divers types de cohérence pouvant intervenir lors de la mesure de quantités fluctuantes est présenté. Des trois types de base de cohérence – temporelle, spectrale et spatiale – l'accent est mis sur la cohérence spatiale du fait que son utilisation en traitement du signal a reçu moins d'attention dans la littérature.

Keywords. Coherence; correlation; temporal coherence; spectral coherence; spatial coherence.

Correspondence to: William A. Gardner, Department of Electrical Engineering and Computer Science, University of California, Davis, CA 95616, USA.

* This paper was first presented in condensed form as the opening plenary lecture at the IEEE Fourth Annual ASSP Workshop on Spectrum Estimation and Modeling, Minneapolis, 3–5 August 1988. This expanded version was prepared with support from the National Science Foundation under Grant MIP88-12902.

1. Introduction

The concept of coherence is fundamental and quite important in all fields dealing with fluctuating quantities. Although there is a commonality among the many uses of the term coherence, the precise meaning of this term seems to vary from one field to another and even within some fields.[1] Some examples of usage include coherent source, coherent field, coherent light, coherent optics, coherent receivers, coherent signals, coherent averaging, coherent processing, temporal coherence, spectral coherence, spatial coherence, self-coherence, mutual coherence, coherence time, coherence bandwidth and so on. The commonality among these uses of the term coherence is captured in the following definition given in the *Oxford English Dictionary*: coherence is "the action or fact of cleaving or sticking together". This dictionary also defines the term correlation as "mutual relation of two or more things". It follows from these definitions that fluctuating quantities that are highly correlated can be said to be coherent. Tying the concept of coherence to that of correlation is helpful because we have precise mathematical definitions of correlation that can be brought to bear on the problem of making precise and quantifying what is meant by coherence.

The purpose of this tutorial paper is to present a unifying view of the concept of coherence that is particularly relevant to the related fields of statistical signal processing and time-series analysis, which permeate various other more specialized fields such as lasers, optics and image processing; communications, radar and telemetry; atmospheric science, oceanography, geophysics and astronomy; sonar, ultrasonics and acoustics; and others. Unfortunately, it is not possible in a single paper to also discuss the various physical origins, historical development and numerous applications to the various fields; but some historical notes are included and a sampling of physical meanings and uses of coherence is presented.

In Section 2, general mathematical definitions of mutual coherence between two fluctuating quantities are given in terms of probabilistic and nonprobabilistic definitions of the correlation coefficient. This is used to establish a link between coherence and statistical linear dependence, and examples are given. This general approach is then applied to special situations that lead to definitions of temporal, spectral and spatial self-coherence, as well as mutual coherence. A spectral decomposition of coherence which leads to the coherency function is then described and used to further develop the link between statistical linear dependence and coherence, and an example is given. The quantification of coherence between two fluctuating quantities is then modified to provide a definition of partial coherence between two fluctuating quantities that are related to other fluctuating quantities. Coherence is then given a geometric interpretation in terms of the angle between two vectors, and the link with statistical linear dependence is described in terms of orthogonal projection. The existence of spectral self-coherence is then tied to the property of cyclostationarity of time-series, and several examples with graphical illustrations are given. The limited utility of spectral coherence for other nonstationarity processes is explained. Finally, the spectral characterization of joint temporal and spectral self-coherence is explained and illustrated graphically with several examples.

In Section 3, a sampling of the many physical meanings of coherence that are commonly used is explained and related to the mathematical definitions presented in Section 2.

In Section 4, some methods for using the various types of coherence that can occur in measurements of fluctuating quantities are briefly described. These include applications of temporal coherence theory of optimum time-invariant filtering for signal extraction and prediction, applications of temporal and spectral coherence theory to frequency-shift filtering for signal extraction and prediction, applications of spatial and spectral coherence

[1] Over 20 years ago, Mandel and Wolf [31] in a comprehensive treatment of optical coherence properties, wrote "No general agreement exists on the precise meaning of the term 'coherence' or the domain encompassed by 'coherence theory'."

theory to antenna array processing for direction-of-arrival estimation, applications of spatial and spectral coherence theory to blind adaptation of antenna arrays for signal extraction (spatial filtering), and applications of temporal and spectral coherence theory to signal detection, classification and time-difference-of-arrival estimation.

Although probabilistic treatments of temporal and/or spatial coherence can be found in many books and papers (cf. [4, 29] for probabilistic treatments of temporal and spatial coherence, respectively), nonprobabilistic treatments and spectral coherence have received relatively little attention [17]. Because of this, the emphasis in this paper is on providing a unifying nonprobabilistic treatment of all three types of coherence with special attention paid to spectral coherence. This emphasis is reflected in the references chosen.

2. Coherence theory

2.1. Degree of coherence

Within a probabilistic framework, the degree of coherence of two zero-mean random variables X and Y is defined to be the magnitude of their correlation coefficient ρ:

$$\rho \triangleq \frac{\mathrm{E}\{XY^*\}}{[\mathrm{E}\{|X|^2\}\mathrm{E}\{|Y|^2\}]^{1/2}} = \frac{R_{XY}}{[R_{XX}R_{YY}]^{1/2}}, \tag{1}$$

where $\mathrm{E}\{\cdot\}$ denotes expecation. For example

$$R_{XY} = \mathrm{E}\{XY^*\} \triangleq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy^* f_{XY}(x, y)\,\mathrm{d}x\,\mathrm{d}y, \tag{2}$$

where f_{XY} is the joint probability density for X and Y. It can be shown (e.g., by using the Cauchy–Schwarz inequality, cf. [15, Section 13.2]) that $0 \leqslant |\rho| \leqslant 1$. For nonzero-mean variables X and Y, the means must be subtracted before applying definition (2). Thus, in general R_{XY} in (1) must be a covariance rather than a correlation.

In practice, given an ensemble of random sample values $X(s)$ and $Y(s)$ with ensemble index s, the empirical degree of coherence can be calculated as follows:

$$|\hat{\rho}| \triangleq \frac{|\hat{R}_{XY}|}{[\hat{R}_{XX}\hat{R}_{YY}]^{1/2}}, \tag{3}$$

where

$$\hat{R}_{XY} \triangleq \frac{1}{N}\sum_{s=1}^{N} X(s)Y^*(s). \tag{4}$$

It can be shown that if $X(s)$ and $Y(s)$ are drawn from a population governed by f_{XY}, then $\hat{R}_{XY}$ converges to R_{XY} in a probabilistic sense (e.g., in mean square or with probability equal to one) as $N \to \infty$.

The degree of coherence is a measure of how closely X and Y are related by a linear transformation. This can be seen from the fact that the minimized fractional mean squared error between X and the linear transformaton aY of Y is given by (cf. [15, Section, 13.3])

$$\min_a \frac{\mathrm{E}\{|X - aY|^2\}}{\mathrm{E}\{|X|^2\}} = 1 - |\rho|^2, \tag{5a}$$

where the minimizing value of a is

$$a = \frac{\mathrm{E}\{XY^*\}}{\mathrm{E}\{|Y|^2\}}. \tag{5b}$$

Thus, X and Y are closely related by a linear transformation if and only if their degree of coherence is close to its maximum value of unity. The two random variables X and Y are said to be completely coherent if and only if $|\rho| = 1$ and completely incoherent if and only if $|\rho| = 0$.

To illustrate the degree of coherence, Fig. 1 shows the scatter plots for three pairs of zero-mean random variables. The scatter plots plot the normalized samples $X'(s) = X(s)/R_{XX}^{1/2}$ and $Y'(s) = Y(s)/R_{YY}^{1/2}$ as the coordinates of points in a plane. In Fig. 1(a), $X \cong aY$ and consequently, $\rho \cong 1$. In Fig. 1(b), X and Y are statistically independent and $\rho = 0$. In Fig. 1(c), $Y \cong X^2$ (where $f_X(\cdot)$ is an even function) and $\rho \ll 1$. This third example illustrates that the degree of coherence is, in general, a measure of the degree of only linear dependence. Variables that are related to a high

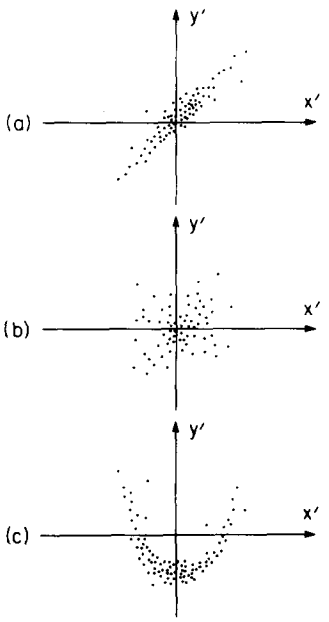

Fig. 1. Scatter plots for three pairs of variables: (a) high degree of coherence; (b) zero degree of coherence; (c) low degree of coherence, but high degree of nonlinear dependence.

degree, but nonlinearly, can have a low degree of coherence. An exception to this is a pair of jointly Gaussian random variables X and Y, which are completely incoherent if and only if they are statistically independent (cf. [15, Section 2.3]). However, if $|\rho| = 1$, then X and Y are completely statistically dependent, regardless of their particular probability distribution.

Within the nonprobabilistic framework of persistent time-series, the degree of coherence of two time-series $x(t)$ and $y(t)$ with zero time-average values is defined to be the magnitude of their temporal correlation coefficient ρ:

$$\rho \triangleq \frac{\langle x(t)y^*(t)\rangle}{[\langle |x(t)|^2\rangle\langle |y(t)|^2\rangle]^{1/2}} = \frac{R_{xy}}{[R_{xx}R_{yy}]^{1/2}}, \tag{6}$$

where $\langle\cdot\rangle$ denotes infinite-time average. For example,

$$R_{xy} = \langle x(t)y^*(t)\rangle \triangleq \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t)y^*(t)\,dt. \tag{7}$$

In practice, given only finite length records of $x(t)$ and $y(t)$, the averaging time T must remain finite. Thus, the empirical degree of coherence can be calculated as follows:

$$|\hat{\rho}| \triangleq \frac{|\hat{R}_{xy}|}{[\hat{R}_{xx}\hat{R}_{yy}]^{1/2}}, \tag{8}$$

where

$$\hat{R}_{xy} \triangleq \frac{1}{T}\int_{-T/2}^{T/2} x(t)y^*(t)\,dt, \tag{9}$$

which converges to R_{xy} as $T\to\infty$. Also, the analogous discrete-time average with time sampling increment δ,

$$\hat{R}_{xy} \triangleq \frac{1}{N}\sum_{n=1}^{N} x(n\delta)y^*(n\delta), \tag{10}$$

can be used in place of (9).

The degree of coherence is a measure of how closely $x(t)$ and $y(t)$ are related by a linear transformation. This follows from the result

$$\min_a \frac{\langle |x(t)-ay(t)|^2\rangle}{\langle |x(t)|^2\rangle} = 1-|\rho|^2, \tag{11a}$$

where

$$a = \frac{\langle x(t)y^*(t)\rangle}{\langle |y(t)|^2\rangle}, \tag{11b}$$

which is the nonprobabilistic counterpart of (5).

The temporal correlation coefficient (6) can be reinterpreted probabilistically as in (1) by introducing the concept of fraction-of-time probability density. Specifically, the fraction-of-time that $x \leqslant x(t) < x+\varepsilon$ and $y \leqslant y(t) < y+\varepsilon$ is given by[2]

$$f_{x(t)y(t)}(x, y; \varepsilon) \triangleq \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} I_{x(t)}(x; \varepsilon)I_{y(t)}(y; \varepsilon)\,dt, \tag{12}$$

[2] Although fraction-of-time probabilities such as (12) are known to exist for time-series that are sample paths from ergodic stochastic processes, it is not necessary to assume that the time-series is associated with a stochastic process for limits such as (12) to exist; see [3, 12].

where $I_{x(t)}(\cdot\,;\varepsilon)$ is the indicator function

$$I_{x(t)}(x;\varepsilon)\triangleq\begin{cases}1, & x\leqslant x(t)<x+\varepsilon,\\ 0, & \text{otherwise,}\end{cases}\tag{13}$$

and similarly for $I_{y(t)}(\cdot\,;\varepsilon)$.The integral in (12) adds up the lengths of the subintervals of $\{-T/2, T/2\}$ over which the joint event $x\leqslant x(t)<x+\varepsilon$ and $y\leqslant y(t)<y+\varepsilon$ occurs. The fraction-of-time probability density of $x(t)$ and $y(t)$ evaluated at x and y is given by the limit

$$f_{x(t)y(t)}(x,y)\triangleq\lim_{\varepsilon\to 0}\frac{1}{\varepsilon^2}f_{x(t)y(t)}(x,y;\varepsilon).\tag{14}$$

In terms of this fraction-of-time density, we have the expected value

$$\mathrm{E}\{x(t)y^*(t)\}\triangleq\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}xy^*f_{x(t)y(t)}(x,y)\,\mathrm{d}x\,\mathrm{d}y.\tag{15}$$

Substitution of (12) and (14) into (15) and interchange of the order of the time-average operation in (12) with the limit in (14) and the two integrals in (15) yields

$$\begin{aligned}&\mathrm{E}\{x(t)y^*(t)\}\\&=\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}\left\{\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_{-\infty}^{\infty}xI_{x(t)}(x;\varepsilon)\,\mathrm{d}x\right.\\&\qquad\left.\times\frac{1}{\varepsilon}\int_{-\infty}^{\infty}y^*I_{y(t)}(y;\varepsilon)\,\mathrm{d}y\right\}\mathrm{d}t\\&=\lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}x(t)y^*(t)\,\mathrm{d}t\\&=\langle x(t)y^*(t)\rangle.\end{aligned}\tag{16}$$

Thus, the time-averages in (6)–(7) can be reinterpreted (if one so chooses) as expected values as in (1)–(2).

In the remainder of this discussion, with one exception, only nonprobabilistic coherence defined in terms of the temporal correlation coefficient (6)–(7) is considered. (The one exception is a brief discussion later in this section on coherence of nonstationary processes.) The reason for emphasizing the nonprobabilistic theory of coherence is that for many applications the conceptual gap between practice and the nonprobabilistic theory is narrower and thus easier to bridge. Moreover, this emphasis complements the emphasis on the probabilistic theory that is found in most other treatments of coherence.[3] (Nevertheless, as suggested by (12)–(16), these dual theories are in fact related by an isomorphism and are, therefore, mathematically equivalent. This is explained in [12, Section 5B and Chapter 15; 15, Section 8.6]).

Since the definition (6) of the degree of coherence $|\rho|$ relates to the correlation between two different fluctuating quantities, it can be called the mutual coherence. This general definition can be specialized in several different ways to define various types of self-coherence involving a single fluctuating quantity. There are three primary specializations of the definition (6) of a degree of coherence for single fluctuating quantities, as explained in the following.

Temporal coherence[4]

Let $x(t)$ and $y(t)$ represent measurements of a single time-series at two different times:

$$x(t)=z(t-t_1),\qquad y(t)=z(t-t_2).\tag{17}$$

In this case,

$$\begin{aligned}R_{xy}&=\langle z(t-t_1)z^*(t-t_2)\rangle\\&=\langle z(t)z^*(t-[t_2-t_1]\rangle\\&=R_{zz}(t_2-t_1),\end{aligned}\tag{18}$$

[3] Although Norbert Wiener, who introduced coherence to the early signal processing community, did use the probabilistic framework of stationary processes, he limited this use to the proof of existence of time series for which the coherence (6) exists and to the proof that averaging in (7) over only $t>0$ or $t<0$ yields the same result; otherwise, he used the nonprobabilistic framework [50].

[4] The study of temporal coherence seems to have its origins in several fields, including optics, economics and meteorology, dating back to the turn of the century (cf. [12, Section 1C]).

which is the autocorrelation function for $z(t)$ evaluated at the time separation $t_2 - t_1$. Similarly,

$$R_{xx} = R_{yy} = R_{zz}(0). \tag{19}$$

Thus, (6) yields

$$\rho = \frac{R_{zz}(\tau)}{R_{zz}(0)}, \tag{20}$$

where $\tau = t_2 - t_1$. This temporal self-coherence can also be generalized to temporal mutual coherence by replacing $x(t)$ and $y(t)$ in (6) with $x(t-t_1)$ and $y(t-t_2)$, respectively.

Spectral coherence

Let $x(t)$ and $y(t)$ represent measurements of a single time-series shifted in frequency by two different amounts:

$$x(t) = z(t)\, e^{-i2\pi f_1 t}, \qquad y(t) = z(t)\, e^{-i2\pi f_2 t}. \tag{21}$$

In this case,

$$R_{xy} = \langle z(t) z^*(t)\, e^{-i2\pi(f_1 - f_2)t} \rangle. \tag{22}$$

More generally, $x(t)$ and $y(t)$ can represent frequency shifted versions of $z(t)$ at two different times,

$$x(t) = z(t-t_1)\, e^{-i2\pi f_1 t}, \qquad y(t) = z(t-t_2)\, e^{-i2\pi f_2 t}, \tag{23}$$

in which case

$$R_{xy} = \langle z(t-t_1) z^*(t-t_2)\, e^{-i2\pi(f_1 - f_2)t} \rangle \tag{24}$$

$$\triangleq R_z^{\alpha}(\tau) \quad \text{for } t_2 = -t_1 = \tau/2 \quad \text{and } \alpha = f_1 - f_2. \tag{25}$$

The function $R_{zz}^{\alpha}(\cdot)$ is called the cyclic autocorrelation function[5] evaluated at cycle frequency α and time separation τ [12]. Since (23) yields

$$R_{xx} = R_{yy} = R_{zz}(0), \tag{26}$$

(6) yields

$$\rho = \frac{R_{zz}^{\alpha}(\tau)}{R_{zz}(0)}, \tag{27}$$

where $\alpha = f_1 - f_2$. This temporal/spectral self-coherence can also be generalized to temporal/spectral mutual coherence by replacing $x(t)$ and $y(t)$ in (6) with $x(t-t_1)\, e^{-i2\pi f_1 t}$ and $y(t-t_2)\, e^{-i2\pi f_2 t}$, respectively. The reason for using notation in which the time-shift τ is an argument within parentheses and the frequency-shift α is a superscript is that we are interested in only individual discrete values of α, whereas we are interested in all values of τ in a continuum. The discreteness of α is explained in Subsection 2.5.

Spatial coherence

Let $x(t)$ and $y(t)$ represent measurements of a single space–time waveform $z(t, \zeta)$ at two different spatial locations ζ_1 and ζ_2:

$$x(t) = z(t, \zeta_1), \qquad y(t) = z(t, \zeta_2). \tag{28}$$

Then the magnitude of (6) yields the degree of spatial self-coherence. More generally, $x(t)$ and $y(t)$ can represent spatial samples of $z(t, \zeta)$ at two different times

$$x(t) = z(t-t_1, \zeta_1), \qquad y(t) = z(t-t_2, \zeta_2). \tag{29}$$

The resultant temporal/spatial self-coherence can also be generalized to temporal/spatial mutual coherence between two space–time waveforms $x(t, \zeta)$ and $y(t, \zeta)$. Furthermore, we can combine the preceding three definitions in an obvious way to obtain the definitions of temporal/spectral/spatial self-coherence and mutual coherence.

2.2. The coherency function

In the cases of temporal/spectral coherence and temporal/spatial coherence, it is possible to spectrally decompose the measure of degree of coherence. This can be accomplished in practice by

[5] The reason for using the adjective cyclic is explained in Subsection 2.5. For finite averaging time T (cf. (7)) and/or for finite energy functions $x(t)$ and $y(t)$, and with deletion of the normalizing factor $1/T$ (cf. (7)), (24) is known as the radar ambiguity function in the very extensive literature on radar, provided that $x(t)$ and $y(t)$ are the analytic signals or complex envelopes corresponding to real signals. But there are important fundamental differences between the theory of cyclic autocorrelation for finite power time-series and the theory of radar ambiguity for finite energy functions, as explained in [12].

passing each of $x(t)$ and $y(t)$ through a tunable narrowband bandpass filter (with transfer function denoted by $H(\cdot)$), and then measuring coherence as a function of the center frequency f of the filter. When this procedure is idealized by letting the bandwidth Δ of the filter approach zero, we obtain the following frequency dependent measure of the degree of coherence:

$$\rho(f) \triangleq \frac{S_{xy}(f)}{[S_{xx}(f)S_{yy}(f)]^{1/2}}, \tag{30}$$

where $S_{xy}(f)$ is the cross spectral density (which is really a spectral correlation density as explained shortly)

$$S_{xy}(f) = \int_{-\infty}^{\infty} R_{xy}(\tau)\, e^{-i2\pi f\tau}\, d\tau, \tag{31}$$

in which $R_{xy}(\tau)$ is the cross correlation function[6]

$$R_{xy}(\tau) \triangleq \langle x(t)y^*(t-\tau)\rangle. \tag{32}$$

For the special case $x(t)=y(t)$, (31) reduces to $S_{xx}(f)$, which is the spectral density of time-averaged power (as explained shortly).

This spectral decomposition is of more limited use for temporal coherence because in this case we have from (17)

$$\begin{aligned} R_{xy}(\tau) &= R_{zz}(\tau + t_2 - t_1), \\ R_{xx}(\tau) &= R_{yy}(\tau) = R_{zz}(\tau). \end{aligned} \tag{33}$$

Thus

$$\rho(f) = e^{i2\pi f(t_2 - t_1)}, \tag{34}$$

and therefore $|\rho(f)| \equiv 1$ for all $x(t)$. Hence, only the phase of $\rho(f)$ can be useful.

To verify that (30) is the correct result, we need to use a few basic results from the theory of stationary time-series [12]. Specifically, for the convolutions

$$y'(t) = y(t) \otimes h_1(t) \triangleq \int_{-\infty}^{\infty} h_1(u)y(t-u)\, du,$$

$$x'(t) = x(t) \otimes h_2(t) \triangleq \int_{-\infty}^{\infty} h_2(u)x(t-u)\, du,$$

we have

$$R_{x'y'}(\tau) = h_2(\tau) \otimes h_1(-\tau) \otimes R_{xy}(\tau),$$

and therefore (using the Fourier transform (31) and the convolution theorem for the Fourier transform)

$$S_{x'y'}(\nu) = H_2(\nu)H_1^*(\nu)S_{xy}(\nu),$$

where

$$H_1(\nu) = \int_{-\infty}^{\infty} h_1(t)\, e^{-i2\pi\nu t}\, dt,$$

and similarly for $H_2(\nu)$; we also have (by inverting the Fourier transform (31))

$$R_{x'y'}(\tau) = \int_{-\infty}^{\infty} S_{x'y'}(\nu)\, e^{i2\pi\nu\tau}\, d\nu. \tag{35}$$

It follows from (7) and (32) that $R_{x'y'} = R_{x'y'}(0)$. Thus we have (using $H_1 = H_2 = H$ in these basic equations)

$$\begin{aligned} R_{x'y'} &= \langle x'(t)y'(t)^*\rangle \\ &= \int_{-\infty}^{\infty} S_{x'y'}(\nu)\, d\nu \\ &= \int_{-\infty}^{\infty} |H(\nu)|^2 S_{xy}(\nu)\, d\nu \\ &= \int_{f-\Delta/2}^{f+\Delta/2} S_{xy}(\nu)\, d\nu \\ &\cong \Delta S_{xy}(f), \end{aligned} \tag{36}$$

[6] In optics the cross correlation function is often called the mutual coherence function and the corresponding correlation coefficient (6) is called the normalized mutual coherence function (cf. [4, 31, 48]).

where f and Δ are the center frequency and bandwidth of the ideal bandpass filter

$$H(\nu)=\begin{cases}1, & |\nu-f|\leq\Delta/2,\\ 0, & \text{otherwise},\end{cases}$$

and Δ is assumed to be small enough to make the approximation (36) accurate.

It follows from (36) that

$$\rho' \triangleq \frac{R_{x'y'}}{[R_{x'x'}R_{y'y'}]^{1/2}} \cong \frac{S_{xy}(f)}{[S_{xx}(f)S_{yy}(f)]^{1/2}}, \tag{37}$$

and the approximation becomes exact as $\Delta \to 0$. This yields the desired result (30).

To see that the cross spectral density (31) is really a spectral correlation density, we simply observe that (36) yields

$$\lim_{\Delta\to 0}\frac{1}{\Delta}R_{x'y'}=S_{xy}(f). \tag{38}$$

Since $R_{x'y'}$ is the correlation of the components of $x(t)$ and $y(t)$ in the spectral band $[f-\Delta/2, f+\Delta/2]$ passed by the filter, then the left member of (38) is obviously a spectral density of correlation evaluated at frequency f. For $x(t)=y(t)$, $R_{x'y'}$ is the time-averaged power (interpreting $x'(t)$ and $y'(t)$ as voltages across one-ohm resistances) of the components of $x(t)$ or $y(t)$ in the spectral band $[f-\Delta/2, f+\Delta/2]$ passed by the filter. Therefore, the left member of (38) with $x(t)=y(t)$ is obviously a spectral density of time-averaged power.

In the statistical literature, the function $\rho(f)$ defined by (30) is typically called the coherency function. Its magnitude can be viewed as a spectrally decomposed measure of the degree to which two time-series are related by a linear time-invariant transformation (a convolution). Specifically, it is shown in [12, Section 7B] that the fractional time-averaged squared error

$$e_{xy} \triangleq \frac{\langle |x(t)-g(t)\otimes y(t)|^2\rangle}{\langle |x(t)|^2\rangle}, \tag{39}$$

when minimized with respect to the impulse-response function $g(\cdot)$, is given by

$$\min_g \{e_{xy}\} = \frac{\int_{-\infty}^{\infty} S_{xx}(f)[1-|\rho(f)|^2]\,df}{\int_{-\infty}^{\infty} S_{xx}(f)\,df}, \tag{40a}$$

where the minimizing function $g(\cdot)$ is given by

$$g(t)=\int_{-\infty}^{\infty}\frac{S_{xy}(f)}{S_{yy}(f)}e^{i2\pi ft}\,df. \tag{40b}$$

In fact, the integrand in the numerator of (40a) is the spectral density of time-averaged power of the optimized error $x(t)-g(t)\otimes y(t)$. This spectral density of error is small at frequency f if and only if the degree of coherence $|\rho(f)|$ between $x(t)$ and $y(t)$ is close to unity at f. Similarly, for e_{yx}, which is obtained by interchanging x and y in the definition of e_{xy}, (39), we obtain (40a) and (40b) with x and y interchanged.

As an example, let us consider a time-series $y(t)$ that is a linearly distorted version $s(t)$ of a signal $x(t)$ plus uncorrelated noise $n(t)$:

$$y(t)=s(t)+n(t), \qquad s(t)=d(t)\otimes x(t),$$

where $d(t)$ is the impulse response function for the distorting transformation. In this case, we find that (cf. [15, Section 13.2])

$$|\rho(f)|=\frac{1}{1+\mathrm{SNR}^{-1}(f)},$$

where

$$\mathrm{SNR}(f)\triangleq\frac{S_{ss}(f)}{S_{nn}(f)}$$

is a frequency-dependent signal-to-noise ratio. Clearly, in frequency bands where $\mathrm{SNR}(f)$ is high (and only in such bands), $y(t)\cong d(t)\otimes x(t)$ is a close approximation and, therefore, $|\rho(f)|\cong 1$ is a close approximation. Thus, $|\rho(f)|$ is seen to be a frequency decomposed measure of the degree to which two time-series are related by a linear time-invariant transformation.

When applied to a single spatial waveform at two points in space, by using (28), the coherency function (30) becomes the self spectral correlation coefficient introduced to optics little more than a decade ago [32]. When applied to a pair of frequency-shifted versions of a single time-series, by using (21), the coherency function (30) becomes the self spectral correlation coefficient introduced to communications engineering only half a decade ago [15, Chapter 12]. This is considered further in Subsection 2.7. Yet the coherency function (30) was in use in time-series analysis as a mutual spectral correlation coefficient at least three decades ago (cf. [34]), and its unnormalized version (31) was introduced by Wiener over four decades ago [50].

2.3. Partial coherence

If we have three time-series $w(t)$, $x(t)$ and $y(t)$, and we want to determine the coherence between $x(t)$ and $y(t)$, with the effects of $w(t)$ removed, we can first subtract from $x(t)$ and $y(t)$ their minimum-time-averaged-squared-error estimates obtained with linear time-invariant transformations of $w(t)$,

$$\begin{aligned} \bar{x}(t) &\triangleq x(t) - g_1(t) \otimes w(t), \\ \bar{y}(t) &\triangleq y(t) - g_2(t) \otimes w(t), \end{aligned} \tag{41}$$

and then obtain the coherence of the residuals $\bar{x}(t)$ and $\bar{y}(t)$. The result,

$$\rho_{\bar{x}\bar{y}}(f) \triangleq \frac{S_{\bar{x}\bar{y}}(f)}{[S_{\bar{x}\bar{x}}(f) S_{\bar{y}\bar{y}}(f)]^{1/2}}, \tag{42}$$

is called the partial coherency function. It is given by [12, Section 7B]

$$\begin{aligned} \rho_{\bar{x}\bar{y}}(f) = & \left[S_{xy}(f) - \frac{S_{xw}(f) S_{yw}^*(f)}{S_{ww}(f)} \right] \\ & \times \left\{ \left[S_{xx}(f) - \frac{|S_{xw}(f)|^2}{S_{ww}(f)} \right] \right. \\ & \left. \times \left[S_{yy}(f) - \frac{|S_{yw}(f)|^2}{S_{ww}(f)} \right] \right\}^{-1/2}. \end{aligned} \tag{43}$$

The same idea can be extended to remove the effects of multiple time-series $\boldsymbol{w}(t) = \{w_1(t), w_2(t), \ldots, w_n(t)\}$ on $x(t)$ and $y(t)$. Specifically, we obtain (42) with [12, Section 7B]

$$\begin{aligned} S_{\bar{x}\bar{y}}(f) = & S_{xy}(f) - \boldsymbol{G}_1'(f) \boldsymbol{S}_{yw}^*(f) \\ & - \boldsymbol{S}_{xw}'(f) \boldsymbol{G}_2^*(f) \\ & + \boldsymbol{G}_1'(f) \boldsymbol{S}_{ww}(f) \boldsymbol{G}_2^*(f), \end{aligned} \tag{44a}$$

$$S_{\bar{y}\bar{y}}(f) = S_{yy}(f) - \boldsymbol{G}_2'(f) \boldsymbol{S}_{ww}(f) \boldsymbol{G}_2^*(f), \tag{44b}$$

$$S_{\bar{x}\bar{x}}(f) = S_{xx}(f) - \boldsymbol{G}_1'(f) \boldsymbol{S}_{ww}(f) \boldsymbol{G}_1^*(f), \tag{44c}$$

where

$$\boldsymbol{G}_1(f) \triangleq [\boldsymbol{S}_{ww}^{-1}(f) \boldsymbol{S}_{wx}(f)]^*, \tag{45a}$$

$$\boldsymbol{G}_2(f) \triangleq [\boldsymbol{S}_{ww}^{-1}(f) \boldsymbol{S}_{wy}(f)]^*. \tag{45b}$$

In these expressions, $\boldsymbol{w}(t)$ is the vector of time-series whose effects on $x(t)$ and $y(t)$ are to be removed, $\boldsymbol{G}_p(f)$ is the vector of transfer functions corresponding to the impulse response vector $\boldsymbol{g}_p(t)$ for $p = 1, 2$ (i.e., $\bar{x}(t) = x(t) - \boldsymbol{g}_1'(t) \otimes \boldsymbol{w}(t)$ as in (41)), $\boldsymbol{S}_{ww}(f)$ is the spectral density matrix with jk-th element $S_{w_j w_k}(f)$, and $\boldsymbol{S}_{wx}(f)$ is the spectral density vector with j-th element $S_{w_j x}(f)$.

2.4. Geometric interpretation of coherence

If we consider for the moment real-valued time-series, then the empirical correlation (10) can be seen to be the N-dimensional Euclidean inner product of the two vectors with elements $\{x(n\delta)\}_1^N$ and $\{y(n\delta)\}_1^N$, scaled by N. In fact, even with the scale factor $1/N$ included and complex-valued time-series allowed, (10) is a valid inner product on a linear vector space. The same is true for the continuous-time counterpart (9) of the discrete-time correlation (10), except that the vectors are now time functions $\{x(t)\}_{-T/2}^{T/2}$ and $\{y(t)\}_{-T/2}^{T/2}$, and the linear vector space is infinite dimensional. Moreover, the same is true of the limit correlation (7). It is a valid inner product of the vectors $\{x(t)\}_{-\infty}^{\infty}$ and $\{y(t)\}_{-\infty}^{\infty}$. This is discussed in more detail in [15, Section 13.2].

The point to be made here is that the numerator of the correlation coefficient (6) is an inner product and its denominator is the product of two norms (the norm is the square root of the inner product of a vector with itself). Consequently, the ratio (6) can be interpreted as the cosine of the angle between the two vectors $\{x(t)\}_{-\infty}^{\infty}$ and $\{y(t)\}_{-\infty}^{\infty}$ (for real time-series). Furthermore, the best linear estimate $ay(t)$ of $x(t)$, which satisfies (11), is the orthogonal projection of the vector $\{x(t)\}_{-\infty}^{\infty}$ onto the one-dimensional linear subspace consisting of all linear transformations $\{ay(t)\}_{-\infty}^{\infty}$ of the vector $\{y(t)\}_{-\infty}^{\infty}$. The closeness of this orthogonally projected vector to the original vector, as indicated by (11), is determined by the closeness of the magnitude of its degree of coherence to unity and, therefore (for real time-series), by the smallness of the angle between the vectors $\{x(t)\}_{-\infty}^{\infty}$ and $\{y(t)\}_{-\infty}^{\infty}$. This geometrical interpretation of coherence is developed in [15, Section 13.3].

Similarly, the best estimate of $x(t)$ obtained by linearly filtering $y(t)$, which satisfies (40), is the orthogonal projection of the vector $\{x(t)\}_{-\infty}^{\infty}$ onto the linear subspace consisting of all linear transformations $\{a_u y(t-u)\}_{t,u=-\infty}^{\infty}$ of the time translates $\{y(t-u)\}_{t,u=-\infty}^{\infty}$ of $\{y(t)\}_{-\infty}^{\infty}$, where $a_u = g(u)$ (cf. [15, Section 13.4]).

In conclusion, the degree of coherence (the magnitude of (6)) as well as its frequency decomposed version (the magnitude of (30)) can always be interpreted (for real time-series) as the magnitude of the cosine of the angle between the two time-series interpreted as vectors. And this degree of coherence is directly related to the accuracy of linear estimation of either time-series in terms of the other, which can be interpreted in terms of the distance between one time-series and its orthogonal projection onto the space spanned by linear transformations of the other.

2.5. *Stationarity and cyclostationarity*

Although mathematical models of time-series, to which the idealized definitions of temporal, spectral and spatial coherence apply, must be stationary[7] in the sense of being persistent so that the infinite-time averages (7) in the definition (6) exist and are finite (neither infinite nor identically zero), we can show that only a very special class of such stationary time-series can exhibit non-zero spectral self-coherence. We can see this from the fact that the cyclic autocorrelation function (25) must not be identically zero for some nonzero cycle frequency α in order to have nonzero spectral coherence (since the numerator of the coherency function (30) is the Fourier transform of (25)), and from the fact that this cyclic autocorrelation function evaluated at some lag value τ is actually the limiting Fourier coefficient (24) of the complex sine-wave component $e^{i2\pi\alpha t}$ contained in the lag-product waveform $z(t+\tau/2)z^*(t-\tau/2)$ for fixed τ. Only time-series $z(t)$ with some underlying periodicity will have such nonzero limiting Fourier coefficients for $\alpha \neq 0$. Furthermore, these Fourier coefficients can be nonzero for only a discrete set of cycle frequencies α [12]. Such time-series are said to exhibit cyclostationarity [12]. If a stationary time-series (one for which $\langle z(t+\tau/2)z^*(t-\tau/2)\rangle$ exists and is finite) exhibits no cyclostationarity, it is said to be purely stationary [12]. For example, a statistical sample of the usual model for thermal noise – an ergodic stationary Gaussian stochastic process – is a purely stationary time-series and, therefore, exhibits no spectral coherence.

Time-series that exhibit cyclostationarity and, therefore, exhibit spectral self-coherence often arise from periodic signal processing operations applied to otherwise purely stationary time-series. Such operations include sampling, scanning, modulating, multiplexing and coding, which arise in communication, telemetry, radar, sonar and control systems. Cyclostationarity can also arise naturally

[7] The class of stationary time-series is broader than the class of all statistical samples of stationary stochastic processes. It includes the statistical samples of asymptotically mean stationary stochastic processes as well. These processes include cyclostationary processes, almost cyclostationary processes, asymptotically stationary processes, and other nonstationary processes (cf. [12, 15]).

from periodicities associated with the physical phenomenon giving rise to the time-series. This includes data obtained from systems subject to seasonal and other rhythmic variations, such as electrocardiograms and other physiological measurements, climatic, oceanic, meteorologic and hydrologic data, and so on.

It should be clarified that periodic components associated with cyclostationarity can arise in the conjugate lag product $z(1+\tau/2)z(t-\tau/2)$ as well as the more commonly used lag product $z(t+\tau/2)z^*(t-\tau/2)$. The corresponding function

$$R^{\alpha}_{zz^*}(\tau) \triangleq \langle z(t+\tau/2)z(t-\tau/2)\, e^{-i2\pi\alpha t}\rangle \qquad (46)$$

is called the conjugate cyclic autocorrelation function.

As an example of a time-series that exhibits cyclostationarity, we consider the amplitude-modulated since wave

$$z(t) = a(t)\cos(\omega_0 t - \theta), \qquad (47)$$

where $a(t)$ is real and purely stationary. Using Euler's formula,

$$\cos(\omega_0 t-\theta) = \tfrac{1}{2}e^{i(\omega_0 t-\theta)} + \tfrac{1}{2}e^{-i(\omega_0 t-\theta)},$$

we obtain

$$\begin{aligned} &z(t+\tau/2)z^*(t-\tau/2)\\ &\quad = \tfrac{1}{4}a(t+\tau/2)a(t-\tau/2)\\ &\qquad \times [e^{i2\omega_0 t}e^{-i2\theta} + e^{-i2\omega_0 t}e^{i2\theta} + e^{i\omega_0\tau} + e^{-i\omega_0\tau}]. \end{aligned}$$

Therefore, we have

$$\langle z(t+\tau/2)z^*(t-\tau/2)\, e^{-i2\pi\alpha t}\rangle$$

$$= \begin{cases} \tfrac{1}{4}\langle a(t+\tau/2)a(t-\tau/2)\rangle e^{\mp i2\theta} & \alpha = \pm\omega_0/\pi,\\ \tfrac{1}{2}\langle a(t+\tau/2)a(t-\tau/2)\rangle \cos(\omega_0\tau), & \alpha = 0,\\ 0, & \text{otherwise,} \end{cases}$$

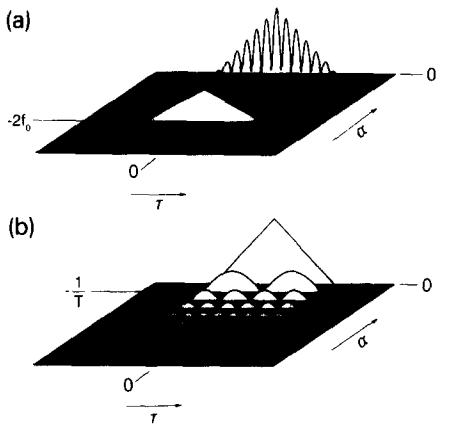

Fig. 2. Graph of temporal/spectral self-coherence function (27) as the height of a surface above the plane with coordinates τ and α: (a) for an amplitude modulated sine wave with carrier frequency f_0; (b) for an amplitude modulated pulse train with pulse rate $1/T$.

because

$$\langle a(t+\tau/2)a(t-\tau/2)\, e^{i2\pi\beta t}\rangle \equiv 0$$

for all $\beta \neq 0$ since $a(t)$ is purely stationary. Thus,

$$R^{\alpha}_{zz}(\tau) = \begin{cases} \tfrac{1}{4}R_{aa}(\tau)\, e^{\mp i2\theta}, & \alpha = \pm\omega_0/\pi,\\ \tfrac{1}{2}R_{aa}(\tau)\cos(\omega_0\tau), & \alpha = 0,\\ 0, & \text{otherwise.} \end{cases} \qquad (48)$$

Hence, $R^{\alpha}_{zz}(\tau)$ is not identically zero for only two nonzero values of α, as shown in Fig. 2(a). The example shown in Fig. 2(a) corresponds to an amplitude time-series $a(t)$ with a triangular autocorrelation.

As another example, we consider the real-valued amplitude-modulated pulse train

$$z(t) = \sum_{n=-\infty}^{\infty} a(nT)p(t-nT), \qquad (49)$$

where $a(t)$ is purely stationary and $p(t)$ is a finite energy pulse,

$$\int_{-\infty}^{\infty} p^2(t)\, dt < \infty.$$

By using the formal characterization

$$z(t)=\left[a(t)\sum_{n=-\infty}^{\infty}\delta(t-nT)\right]\otimes p(t)$$

$$=\left[a(t)\frac{1}{T}\sum_{n=-\infty}^{\infty}e^{i2\pi mt/T}\right]\otimes p(t),$$

$$=w(t)\otimes\frac{1}{T}p(t),$$

where

$$w(t)=a(t)\sum_{m=-\infty}^{\infty}e^{i2\pi mt/T},$$

we can show that

$$R_{zz}^{\alpha}(\tau)=\frac{1}{T^2}R_{ww}^{\alpha}(\tau)\otimes r_p^{\alpha}(\tau), \tag{50}$$

where

$$r_p^{\alpha}(\tau)\triangleq\int_{-\infty}^{\infty}p(t+\tau/2)p(t-\tau/2)\,e^{i2\pi\alpha t}\,dt.$$

By an argument similar to that made for the amplitude-modulated sine wave example, we can show that

$$R_{ww}^{\alpha}(\tau)=R_{aa}(\tau)\sum_{m=-\infty}^{\infty}e^{-i2\pi m\tau/T}$$

$$=R_{aa}(\tau)T\sum_{m=-\infty}^{\infty}\delta(\tau-nT),\quad \alpha=q/T$$

for all integers q. Therefore, (50) yields

$$R_z^{\alpha}(\tau)=\frac{1}{T}\sum_{n=-\infty}^{\infty}R_{aa}(nT)r_p^{\alpha}(\tau-nT),$$

$$\alpha=q/T \tag{51}$$

for all integers q. Thus, $R_{zz}^{\alpha}(\tau)$ is not identically zero for only values of α that are integer multiples of $1/T$, as shown in Fig. 2(b). The example shown in Fig. 2(b) corresponds to a white amplitude sequence $a(nT)$ and a rectangular pulse $p(t)$ of width T.

2.6. Nonstationarity

It is interesting that within a probabilistic framework, we can define spectral self-coherence for nonstationary stochastic processes other than those that exhibit cyclostationarity. All that is required is that the probabilistic autocorrelation

$$\underset{\sim}{R}_{XX}(t,\tau)\triangleq E\{X(t+\tau/2)X^*(t-\tau/2)\} \tag{52}$$

depends on t and does not contain any periodic components in t (other than a constant component):

$$\langle\underset{\sim}{R}_{XX}(t,\tau)\,e^{-i2\pi\alpha t}\rangle\equiv 0,\quad \alpha\neq 0.$$

All such nonstationary stochastic processes exhibit probabilistic spectral self-coherence but not cyclostationarity. However, unlike the spectral self-coherence associated with cyclostationarity, it is a property of an ensemble of time-series, not a single time-series, and it therefore cannot in general be reliably estimated (and hence cannot be used in practice) without using ensemble averages instead of time averages [13]. Unfortunately, ensembles from nonstationary stochastic processes are rarely available in practice. There are two (and only two) exceptions to this generality as explained in [13]. If either the nonstationarity (which is assumed to be not cyclostationarity) is sufficiently slow that the process is locally stationary (nearly stationary over relatively long periods) and locally ergodic[8] or the form of nonstationarity is known (e.g., $\underset{\sim}{R}_{XX}(t,\tau)$ is a combination of functions of t alone and functions of τ alone, and the functions of t alone are known), then (and only then) the autocorrelation and therefore the temporal and spectral coherence can be reliably estimated using only time-averaging operations on a single time-series.

It is also interesting that all spectral self-coherence in nonstationary stochastic processes that is not associated with cyclostationarity must be zero on the average over time. This can be seen as follows. To obtain a measure of the correlation between spectral components in a band of width Δ

[8] The subtleties of locally ergodic probabilistic models are discussed in [12].

centered at frequency $f+\alpha/2$ and another band of width Δ centered at frequency $f-\alpha/2$, we simply pass the process through a pair of corresponding bandpass filters to obtain the two processes

$$X_+(t)=h_+(t)\otimes X(t),$$
$$X_-(t)=h_-(t)\otimes X(t), \tag{53}$$

then we frequency-shift these to a common band, e.g.,

$$Y_+(t)=X_+(t)\,e^{-i2\pi(f+\alpha/2)t},$$
$$Y_-(t)=X_-(t)\,e^{-i2\pi(f-\alpha/2)t}, \tag{54}$$

and finally we obtain their correlation

$$\underline{R}_\alpha=E\{Y_+(t)Y_-^*(t)\}. \tag{55}$$

Substitution of (53) into (54) and the result into (55) and interchange of the order of operations of expectation and integration yields

$$\underline{R}_\alpha=\iint h_+(u)h_-^*(v)\underline{R}_{XX}\left(t-\frac{u+v}{2},\,v-u\right)\times e^{-i2\pi\alpha t}\,du\,dv. \tag{56}$$

The average of this spectral correlation over time t will be nonzero if and only if

$$\langle\underline{R}_{XX}(t,\tau)\,e^{-i2\pi\alpha t}\rangle\not\equiv 0,\quad \alpha\neq 0,$$

which would indicate that $X(t)$ exhibits cyclostationarity with cycle frequency α.

The presence of the factor $e^{-i2\pi\alpha t}$ in this average is a direct result of the down conversion operation (54). However, if this operation is eliminated, then we obtain

$$\underline{R}_\alpha=\iint h_+(u)h_-^*(v)\underline{R}_{XX}\left(t-\frac{u+v}{2},\,v-u\right)du\,dv$$
$$=\iint H_+(\mu)H_-^*(\nu)\underline{S}_{XX}^*\left(\mu-\nu,\frac{\mu+\nu}{2}\right)\times e^{-i2\pi(\nu-\mu)t}\,d\mu\,d\nu, \tag{57b}$$

where

$$\underline{S}_{XX}(f,g)\triangleq\iint\underline{R}_{XX}(t,\tau)\,e^{-i2\pi(ft-g\tau)}\,dt\,d\tau$$

and

$$H_\pm(f)\triangleq\int h_\pm(t)\,e^{-i2\pi ft}\,dt.$$

The average of this correlation (57b) over time t can be nonzero only if the integrand is not identically zero for $\nu=\mu$. But as long as the passbands of the two bandpass filters with center frequencies $f+\alpha/2$ and $f-\alpha/2$ do not overlap, then the product of their transfer functions $H_+(\mu)H_-^*(\mu)$ must be identically zero and therefore the integrand is identically zero (even if $X(t)$ exhibits cyclostationarity).

In summary, the concept of spectral self-coherence is of limited utility (in the sense that it is not a property exhibited by a single time-series, such as a statistical sample of a stochastic process) for nonstationary stochastic processes that do not exhibit cyclostationarity and are not locally stationary. Examples of the former case are described in the preceding subsection (and following subsection). An example of the latter case is given here.

A process $X(t)$ is said to be locally stationary (in the wide sense) if the autocorrelation function (52) is nearly constant in the location variable t over all intervals of length T for some T that is much larger than the width of the autocorrelation function (52) in the lag variable τ. We consider the extreme case of a locally stationary process for which the probabilistic autocorrelation is given by

$$\underline{R}_{XX}(t,\tau)=a^2(t)\delta(\tau),$$

where $\delta(\tau)$ is the Dirac delta function. This autocorrelation corresponds to the process $X(t)=a(t)N(t)$, where $N(t)$ is white noise. It follows that (56) reduces in this case to

$$\underline{R}_\alpha=[g(t)\otimes a^2(t)]\,e^{-i2\pi\alpha t},$$

where $g(t)$ is the inverse Fourier transform of

$$G(f)\triangleq H_+(f)\otimes H_-^*(-f),$$

which is the transfer function of a bandpass filter

with center frequency α and bandwidth 2Δ. Thus, the stronger the spectral content of $a^2(t)$ in the vicinity of frequency α, the stronger the spectral correlation of $X(t)$ for all frequencies separated by α. The average of this spectral correlation $\underset{\sim}{R}_{XX}$ over all time t is, however, zero unless $a^2(t)$ contains a finite-strength additive sine-wave component with frequency α, in which case $X(t)$ exhibits cyclostationarity.

This idealized example suggests that, in general, the stronger the degree of spectral self-coherence that can be reliably estimated from a single time-series, the stronger the degree of local cyclostationarity. In fact, it is shown in [18] that for a locally stationary process the spectral correlation coefficient magnitude for widely separated bands must be much less than 1/2 when the spectral content of $a(t)$ is flat, but can be as large as 1/2 when the spectral content of $a(t)$ is highly peaked at $\alpha/2$.

In contrast to the relatively limited utility of the concept of spectral self-coherence for nonstationary stochastic processes, the concepts of temporal and spatial self-coherence are generally useful for nonstationary processes. Although these time-variant (nonstationary) degrees of coherence can be accurately measured only with the use of ensemble averaging (except for locally stationary processes), their time-averaged values can be accurately measured using only time averaging on a single (pair of) time-series [single statistical sample of the (pair of) stochastic process(es)] (cf. [12, Section 8.5; 13]). For example, the time-averaged temporal self-coherence

$$\rho \triangleq \frac{\langle \underset{\sim}{R}_{XX}(t, \tau) \rangle}{\langle \underset{\sim}{R}_{XX}(t, 0) \rangle},$$

where $\langle \cdot \rangle$ denotes averaging over the time parameter t, can be accurately estimated using (8) and (17), where $z(t)$ is a statistical sample of $X(t)$.

2.7. *Spectral characterization of temporal and spectral coherence*

Since the temporal coherence of a time-series $z(t)$ is completely determined by its autocorrelation function (20), then it can be characterized by the Fourier transform of its autocorrelation,

$$S_{zz}(f) = \int_{-\infty}^{\infty} R_{zz}(\tau)\, e^{-i2\pi f\tau}\, d\tau, \tag{58}$$

which (from (31) and (38)) is the spectral density of time-averaged power in $z(t)$.

The power spectral density $S_{zz}(f)$ has long been used as a descriptor of stationary time-series $z(t)$. But for time-series that exhibit cyclostationarity, there is another more general descriptor that characterizes both the temporal and spectral coherence properties of the time-series. This descriptor is, by analogy with (58), the Fourier transform of the cyclic autocorrelation function (25) (sometimes called the cyclic spectral density):

$$S_{zz}^{\alpha}(f) \triangleq \int_{-\infty}^{\infty} R_{zz}^{\alpha}(\tau)\, e^{-i2\pi f\tau}\, d\tau. \tag{59}$$

It follows directly from (24)–(25), (38) and (59) that, with $\alpha = f_1 - f_2$,

$$S_z^{\alpha}(f) = \lim_{\Delta \to 0} \frac{1}{\Delta} R_{x'y'}, \tag{60}$$

where $R_{x'y'}$ is the correlation of the components of $x(t)$ and $y(t)$ in the spectral band $[f-\Delta/2, f+\Delta/2]$ which, from (23), is the correlation of the time shifted components of $z(t)$ in the two spectral bands $[f+f_1-\Delta/2, f+f_1+\Delta/2]$ and $[f+f_2-\Delta/2, f+f_2+\Delta/2]$. Thus, $S_{zz}^{\alpha}(f)$ is the density of correlation between the spectral components in $z(t)$ at frequencies separated by $\alpha = f_1 - f_2$. For $\alpha = 0$, this reduces to the spectral density of time-averaged power in $z(t)$.

This spectral correlation function can be made into a spectral correlation coefficient as in (30). It follows from (23) that the factors in the denominator of (30) are given by

$$S_{xx}(f) = S_{zz}(f+f_1), \qquad S_{yy}(f) = S_{zz}(f+f_2). \tag{61}$$

It follows from (25), (30), (59) and (61) that this spectral correlation coefficient (the self spectral

coherency function) is given by

$$\rho^{\alpha}(f)=\frac{S_{zz}^{\alpha}(f)}{[S_{zz}(f+\alpha/2)S_{zz}(f-\alpha/2)]^{1/2}},$$

$$\text{for } f_1=-f_2=\alpha/2. \quad (62)$$

Since a time-series $z(t)$ can exhibit cyclostationarity at only a discrete set of cycle frequencies α [12], the spectral correlation function $S_{zz}^{\alpha}(f)$ is discrete in α, although it is generally continuous in f. This spectral characterization of the temporal and spectral coherence properties of time-series has been calculated and graphed in [12] for a wide variety of different types of modulated signals. Two examples are presented here.

For the first example, we consider a sine wave that is amplitude-modulated by a purely stationary time-series as in (47). It follows from the formula (48) for the cyclic autocorrelation for this time-series and the Fourier transform relation (59) that the spectral correlation function is given by

$$S_{zz}^{\alpha}(f)=\begin{cases}\frac{1}{4}S_{aa}(f)\,e^{\mp i2\theta}, & \alpha=\pm 2f_0,\\ \frac{1}{4}S_{aa}(f+f_0)+\frac{1}{4}S_{aa}(f-f_0), & \alpha=0,\\ 0, & \text{otherwise,}\end{cases} \quad (63)$$

where $f_0=\omega_0/2\pi$. Thus, only spectral components that are separated by $|\alpha|=2f_0$ are correlated. This can be easily understood by expressing the time-series (47) as

$$z(t)=\tfrac{1}{2}a(t)\,e^{i2\pi f_0 t}\,e^{-i\theta}+\tfrac{1}{2}a(t)\,e^{-i2\pi f_0 t}\,e^{i\theta}.$$

Thus, each spectral component in $a(t)$ is shifted from its original frequency, say f, to $f+f_0$ and $f-f_0$. The separation between the frequencies of these pairs of identical (except for a constant phase difference of 2θ) spectral components is $2f_0$. Hence, all such components are completely correlated (as long as $S_{aa}(f)=0$ for $|f|\geq f_0$ so that no positively shifted components overlap with any negatively shifted components). That is, it follows from (63) that $|\rho^{\alpha}(f)|=1$ for $\alpha=\pm 2f_0$. A graph of $|S_{zz}^{\alpha}(f)|$ interpreted as the height of a surface above the plane with coordinates f and α is shown in Fig. 3(a) for the same example as that in Subsection 2.5. Notice that since $S_{zz}^{\alpha}(f)$ must be discrete in α, such surfaces will always consist of infinitesimally thin slices parallel to the f axis. Also, the slice along $\alpha=0$ is the power spectral density function.

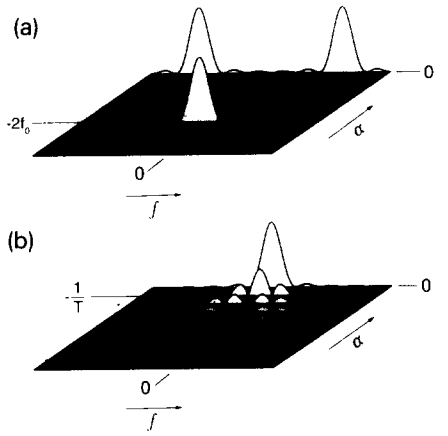

Fig. 3. Theoretical spectral correlation magnitudes (a) for an amplitude-modulated sine wave with carrier frequency f_0; (b) for an amplitude-modulated pulse train with pulse rate $1/T$.

For the second example, we consider a periodic pulse train that is amplitude-modulated by a purely stationary discrete-time time-series as in (49). It follows from (51) for the cyclic autocorrelation for this time-series and the Fourier transform relation (59) that the spectral correlation function is given by (cf. [12])

$$S_{zz}^{\alpha}(f)=\frac{1}{T^2}P(f+\alpha/2)P^{*}(f-\alpha/2)\times\sum_{n=-\infty}^{\infty}S_{aa}(f-\alpha/2-n/T),\quad \alpha=k/T \quad (64)$$

for all integers k, and it is zero for all other values of α. In (64), $P(f)$ is the Fourier transform of the pulse $p(t)$ in (49). Thus, only spectral components that are separated by integer multiples of $1/T$ are correlated. This can be easily understood by expressing the time-series (49) as

$$z(t)=w(t)\otimes\frac{1}{T}p(t),$$

where

$$w(t)=a(t)\sum_{k=-\infty}^{\infty} e^{i2\pi kt/T}.$$

Thus, each spectral component in $a(t)$ is shifted by all integer multiples k of $1/T$. As long as $S_{aa}(f)=0$ for $|f|>1/2T$ so that no shifted components overlap with each other, then all spectral components are perfectly correlated. That is, it follows from (64) that $|\rho^{\alpha}(f)|=1$ for $\alpha=k/T$ for all k for which the power density at $f\pm\alpha/2$ is nonzero: $S_{zz}(f\pm k/2T)\neq 0$. A graph of the spectral correlation surface $|S_{zz}^{\alpha}(f)|$ for a flat spectrum $S_{aa}(f)$ ($S_{aa}(f)=$constant for $|f|\leq 1/2T$ and $S_{aa}(f)=0$ for $|f|>1/2T$) and a flat pulse $p(t)$ ($p(t)=0$ for $|t|>T/2$), which is the same as the example in Subsection 2.5, is shown in Fig. 3(b).

Other examples including sine waves that are phase- or frequency-modulated by stationary time-series as well as by amplitude-modulated periodic pulse trains, and also periodic pulse trains that are pulse-width or pulse-position modulated are presented in [12].

3. Some physical meanings of coherence

3.1. Coherent electromagnetic fields

The distinction between coherent and incoherent radiation, say light, is not clear cut [5], and also is not based solely on the mathematical definitions of coherence given in Section 2. For example, a narrowband wave of light observed at one point in space,

$$z(t)=a(t)\cos[\omega_0 t+\phi(t)],$$

obtained by passing incoherent incandescent radiation (white light) through some narrowband optical filter, exhibits a high degree of temporal coherence in both its envelope $a(t)$ and phase $\phi(t)$, in the sense that

$$\rho_a=\frac{R_{aa}(\tau)}{R_{aa}(0)} \quad \text{and} \quad \rho_{\phi}=\frac{R_{\phi\phi}(\tau)}{R_{\phi\phi}(0)},$$

are very close to unity for relatively large values of time separation τ; that is, values up to the coherence time, which is approximately the reciprocal of the bandwidth of $a(t)$ and $\phi(t)$ which is equal to the bandwidth of the filter. The longer the coherence time is relative to the period corresponding to the center frequency of the filter, ω_0, the higher the degree of monochromaticity (which is inversely related to the relative bandwidth) will be. If this spectrally filtered light is then spatially filtered and colimated, it also exhibits a high degree of spatial coherence. However, it is more practical to obtain high degrees of temporal and spatial coherence over relatively long times (much greater than the propagation time across an optical system) and relatively wide beams by using light amplification by stimulated emission – a laser [47]. Moreover, there is a distinction between laser light and filtered incandescent light that cannot be explained in terms of our mathematical definitions of coherence. Specifically, laser light has a highly stable envelope $a(t)$ with relatively small variation about a positive mean value. As a result, the joint fraction-of-time density of the in-phase and quadrature components $c(t)$ and $s(t)$,

$$z(t)=c(t)\cos(\omega_0 t)-s(t)\sin(\omega_0 t),$$

where

$$c(t)=a(t)\cos\phi(t), \qquad s(t)=a(t)\sin\phi(t),$$

looks like a continuous mountain range that forms a ring centered at the origin of the plane. In contrast to this, the joint fraction-of-time density for $c(t)$ and $s(t)$ for filtered incandescent light looks like a single mountain peaked at the origin. For both types of light, the phase $\phi(t)$ has a uniform fraction-of-time density throughout the interval $[-\pi, \pi]$.

Generalizing on the preceding, a spatially extended source of radiation is called a coherent source if the signals (say (28)) emitted at different points on the source exhibit a sufficiently high degree of mutual temporal coherence (6). Otherwise it is said to be an incoherent source [48]. Also, if we let $u(t, \zeta)$ denote the complex envelope of

the space-time waveform $x(t, \zeta)$ (i.e., the positive-frequency portion of $x(t, \zeta)$ shifted in frequency so that its center frequency is near zero), then the time averaged value of $\langle u(t, \zeta)\rangle$ is typically called the coherent field, and the residual $u(t, \zeta) - \langle u(t, \zeta)\rangle$ is called the incoherent field [28]. Thus, the coherent field is zero for incoherent light as defined above. Similarly, the time averaged squared magnitudes of these two quantities are called the coherent field intensity and the incoherent field intensity, respectively [28].

Attempts to generalize coherence from a measure of linear dependence to a measure of nonlinear dependence have been made, for example, in the field of quantum optics [23]. This has led to the use of the term coherent states not only in quantum optics [24], but also in signal theory [11]. However, these coherent states in signal theory are nothing more than an over-complete set of basis functions that can be used for the representation of signals in the joint time/frequency domain [27]. There is no direct connection between the meaning of coherence discussed in this paper and its (perhaps unfortunate) use in the term coherent states for particular sets of basis functions.

3.2. Coherence time and bandwidth of propagation media

When a single-frequency signal – a sine wave – propagates through a randomly time-varying medium such as the ocean or the ionosphere, the received signal will be spread in frequency. As a result, its temporal self-coherence function (20) will have a finite width in τ (whereas the coherence function of the transmitted sine wave is itself a sine wave that extends forever in τ), which is on the order of the reciprocal of the width of the power spectral density function. The width of the temporal self-coherence function is called the coherence time of the medium [28].

When two sine-waves with distinct frequencies f_1 and f_2 propagate simultaneously through a randomly time-varying medium, and the two corresponding received signals are each frequency-shifted to zero frequency (nominally) and their mutual coherence (6) is measured, the result will decrease from $|\rho| = 1$ as $|f_1 - f_2|$ increases until eventually $|\rho| \ll 1$. That is, the mutual spectral coherence[9] (analogous to the spectral self-coherence (27)) will exhibit a finite width in the parameter $\alpha = f_1 - f_2$. This width is called the coherence bandwidth of the medium [28].

3.3. Coherent signals in electrical circuits

The distinction between what are sometimes called coherent and incoherent signals that are found, for example, in radio and radar circuits is quite different from the distinction for light waves. Whereas the latter distinction involves the envelope, the former distinction is based primarily on the phase. For example, a sequence of sine wave voltage bursts, such as in an amplitude-shift-keyed signal (i.e., a sine wave whose amplitude is periodically shifted among a finite set of values) can all be in phase with each other or their phases can change randomly from one burst (keying interval) to the next. The former type of signal is said to be coherent, whereas the latter is called incoherent. Thus, a coherent radio signal can have large fluctuations in its envelope, but its phase must be almost time-invariant.

On the other hand, the concept of a stable envelope can be used to distinguish between two types of narrowband radio signals that also are commonly labeled coherent and incoherent. Specifically, the output of an electrical oscillator, like laser light, typically has a highly stable envelope but a phase that fluctuates randomly throughout the interval $[-\pi, \pi]$, although this fluctuation is relatively slow (depending on the effective bandwidth of the oscillator) – the width of the temporal self-coherence function of the phase is called the coherence time of the oscillator. In contrast to this, the output of a narrowband bandpass filter

[9] Notice that for a stationary random channel, the pair of received signals are jointly cyclostationary with cycle frequency $\alpha = f_1 - f_2$ (as long as $|f_1 - f_2|$ does not exceed the coherence bandwidth of the channel), in which case the signals are uncorrelated with each other.

with the same bandwidth as the effective bandwidth of the oscillator and with a broadband noise input will have an envelope that fluctuates randomly, completely analogously to the filtered incoherent light described in the previous subsection.

Nevertheless, any narrowband signal is often called coherent if its amplitude and phase are essentially constant throughout whatever time interval is of interest for purposes of detection, measurement, processing and so on, regardless of what their fraction-of-time densities might be over time-intervals that are longer, e.g., long enough to obtain statistically reliable measurements of fraction-of-time densities. But this is consistent with the mathematical definition of temporal coherence since (63) and (64) would both be close to unity for all time-separations τ of interest (which we have assumed are less than the reciprocal bandwidth).

In addition to these temporal coherence properties of radio signals, essentially all modulated signals regardless of bandwidth, by virtue of their cyclostationarity, exhibit spectral self-coherence properties as explained and illustrated in Section 2. In contrast to this, purely stationary noise exhibits no spectral self-coherence, and purely random (white) noise exhibits no temporal self-coherence since

$$R_{zz}(\tau) = \int_{-\infty}^{\infty} S_{zz}(f)\, e^{i2\pi f\tau}\, df$$

$$= N_0\delta(\tau) = 0, \quad \tau \neq 0,$$

for the flat power spectral density of white noise

$$S_{zz}(f) = N_0.$$

3.4. *Coherent signal processors*

The term coherent, as applied to an optical processor, a radio communication receiver, a radar receiver, or other signal processor, means that the instantaneous phase of narrowband signals is kept track of and utilized for some purpose. For example, in coherent radio, the phase of a locally generated sine wave is locked onto the phase of the sine-wave carrier in the received signal, so that the signal can be coherently demodulated. In coherent radar, the phase of the received signal relative to that of the transmitted signal is measured and utilized. In adaptive coherent optical processing, a received distorted wavefront is phase-compensated so that all points on the wavefront are in phase. In contrast to this, the phase is ignored by noncoherent processors. (The term noncoherent is often used here in place of the term incoherent. This might be due to the common use of incoherent to mean unintelligible, e.g., expression of thoughts in sequence that do not cohere with each other.) In signal processing in general, the term coherent averaging is used to mean averaging of a phase-dependent quantity, such as a complex envelope, whereas noncoherent averaging denotes averaging of a phase-independent quantity such as the magnitude of a complex envelope.

4. Some uses of coherence

In the allied fields of signal processing and time-series analysis, temporal, spectral and spatial coherences, as defined in Section 2, are typically useful because they correspond directly to the accuracy with which one fluctuating quantity can be approximated or estimated by a linear transformation of another fluctuating quantity.[10] Some examples of this are mentioned here. Because of limitations on the length of this paper, only brief descriptions of these examples are given. However, references to more detailed treatments (containing extensive reference lists) are given. Also, because

[10] In spite of this fact, it has been said that the role of coherence in signal processing is often not explicitly recognized. For example, it is stated in [7] that "The coherence function provides a mechanism for measuring the linear association existent between two time series. It has been used by statisticians and scientists alike for this purpose, but its utilization by the signal processing community is somewhat limited." If this is true, it is a curious state of affairs, since half a century has passed since Norbert Wiener – who is credited with initiating the confluence of the fields of time-series analysis and communication engineering, which then evolved into statistical signal processing – introduced coherence to the early signal processing community (cf. [50]).

of the relative novelty of methods for exploiting spectral coherence, a little more emphasis is put on examples of the usage of this particular type of coherence.

Although the methods of exploiting coherence that are surveyed here are most likely to be implemented with digital signal processing algorithms, the methods are described in terms of continuous-time signal processing operations because of their closer ties with the underlying physics and the corresponding continuous-time theory discussed in preceding sections. Typically, the translation of methods from continuous time to discrete time is straightforward.

4.1. *Wiener filtering: temporal mutual coherence*

When we have two distinct but partially coherent time-series $x(t)$ and $y(t)$, we can use the temporal coherence between them to estimate one, say $x(t)$, by filtering the other, $y(t)$, that is by adding up linearly transformed versions of time-shifted replicas of $y(t)$. The transfer function

$$G(f)=\int_{-\infty}^{\infty} g(t)\, e^{-i2\pi ft}\, dt$$

of the filter that minimizes the time-averaged squared error

$$\langle |x(t)-\hat{x}(t)|^2 \rangle \tag{65}$$

between $x(t)$ and its estimate

$$\hat{x}(t) \triangleq \int_{-\infty}^{\infty} g(u)y(t-u)\, du$$

is given by [12, Section 7b]

$$G(f)=\frac{S_{xy}(f)}{S_{yy}(f)}, \tag{66}$$

which is known as the noncausal Wiener filter [50]. For example, if $x(t)$ and $y(t)$ are related by

$$y(t)=s(t)+n(t),$$

where

$$s(t)=x(t)\otimes d(t),$$

and $d(t)$ represents signal distortion and $n(t)$ represents zero-mean additive noise that is uncorrelated with $s(t)$, then (66) reduces to

$$G(f)=\frac{1}{D(f)}|\rho(f)|^2,$$

where $\rho(f)$ is the coherency function

$$\rho(f) \triangleq \frac{S_{xy}(f)}{[S_{xx}(f)S_{yy}(f)]^{1/2}}=\frac{1}{1+\dfrac{S_{nn}(f)}{S_{ss}(f)}}.$$

Thus, at frequencies f where the signal-to-noise ratio (SNR) $S_{ss}(f)/S_{nn}(f)$ is high, the coherency magnitude is close to unity and the filter $G(f)$ simply removes the distortion by using the inverse of the distortion operation. However, at frequencies where the SNR is low, the coherency magnitude is much less than unity and the filter $G(f)$ acts primarily as an attenuator to suppress these noisy spectral components.

As another example, if $x(t)$ and $y(t)$ are related by

$$x(t)=s(t)+n(t), \qquad y(t)=m(t),$$

where $s(t)$ is a signal, $n(t)$ is additive noise and $m(t)$ represents a measurement of noise only that is correlated with $n(t)$, then the temporal coherence between $n(t)$ and $m(t)$ can be used to partially cancel the noise $n(t)$ that corrupts the signal. If $m(t)$ has zero mean and is uncorrelated with $s(t)$, then it follows from (66) that the optimum noise-estimation filter has transfer function

$$G(f)=\frac{S_{nm}(f)}{S_{mm}(f)}.$$

The residual noise remaining after the noise estimate

$$\hat{n}(t)=g(t)\otimes y(t)$$

is subtracted from $x(t)$ has fractional time-averaged squared value

$$\frac{\int_{-\infty}^{\infty} S_{nn}(f)[1-|\rho(f)|^2]\,df}{\int_{-\infty}^{\infty} S_{nn}(f)\,df},$$

where $\rho(f)$ is the coherency function for $n(t)$ and $m(t)$. Thus, the extent to which $n(t)$ can be cancelled in each frequency band of interest is determined by the coherency function.

For more detailed treatments see [15, Chapter 13] and the many references therein.

4.2. *Wiener prediction: temporal self-coherence*

If we have one time series $x(t)$ and we want to predict its future $x(t+t_0)$ using linear time-invariant transformations on its own past $\{x(u): u \leqslant t\}$, the degree of accuracy of the predictions, as measured by the fractional time-averaged squared prediction error, is completely determined by the degrees of temporal coherence of $x(t)$ for all possible time shifts:

$$\min_{g(\cdot)} \frac{\langle |x(t+t_0)-\hat{x}(t+t_0)|^2\rangle}{\langle |x(t)|^2\rangle} = 1-\int_0^{\infty} g(\tau)\rho(\tau+t_0)\,d\tau,$$

where

$$\hat{x}(t+t_0)=\int_{-\infty}^{t} g(t-u)y(u)\,du,$$

and the impulse-response function $g(u)$ of the optimum prediction filter is the solution to the Wiener–Hopf equation [50; 15, Chapter 13]

$$\int_0^{\infty} g(u)\rho(\tau-u)\,du=\rho(\tau+t_0), \quad \tau\geqslant 0,$$

where $\rho(\tau)=R_{xx}(\tau)/R_{xx}(0)$. The greater the degree of temporal coherence $\rho(\tau)$ for $\tau\geqslant\tau_0$, the more accurate the predictions will be. For more detailed treatments see [12, 15] and the many references therein.

4.3. *Cyclic Wiener filtering and prediction: temporal and spectral coherence*

If the time series $x(t)$ and $y(t)$ in either of the filtering and prediction problems just described exhibit cyclostationarity, then they possess spectral coherence as well as temporal coherence. Thus, frequency-shifted as well as time-shifted versions of one of the time-series (or its past) can be linearly combined to form an estimate of the other time-series for the filtering problem (or of the future of the same time-series for the prediction problem). For example, if the cyclic correlation functions $R_{xy}^{\alpha}(\tau)$ and $R_{yy}^{\beta}(\tau)$ are nonzero for some nonzero values of α and β, then the smallest value of estimation error (65) obtainable using only linear transformations can be obtained using an estimate of the form

$$\hat{x}(t)=\sum_{\gamma}\int_{-\infty}^{\infty} g_{\gamma}(u)y(t-u)\,e^{-i2\pi\gamma(t-u)}\,du,$$

where the sum ranges over all the difference frequencies $\gamma=\alpha-\beta$ [12, Section 14A; 15, Section 12.8]. This estimate, which is a sum of linearly transformed (with time-invariant impulse-response functions $\{g_{\gamma}(\cdot)\}$) frequency-shifted versions of $y(t)$, can be reexpressed as the output

$$\hat{x}(t)=\int_{-\infty}^{\infty} h(t,v)y(v)\,dv$$

of a single linear transformation with multiply-periodically-time-variant impulse-response function

$$h(t,v)=\sum_{\gamma} g_{\gamma}(t-v)\,e^{-i2\pi\gamma v}.$$

The transfer functions $\{G_{\gamma}(f)\}$ corresponding to the optimum impulse-response functions $\{g_{\gamma}(t)\}$ (those that minimize the estimation error (65)) are the solutions to the set of simultaneous linear

equations

$$\sum_{\gamma} G_{\gamma}(f) S_{yy}^{\alpha-\gamma}\left(f-\frac{\alpha+\gamma}{2}\right)=S_{xy}^{\alpha}(f-\alpha/2)$$

for all values of f and all values of α for which the spectral correlation functions are not identically zero. This optimum multiply-period filter is called the cyclic Wiener filter [19]. The value of the minimized estimation error (65) is given by [12, Section 14A; 15, Section 12.8]

$$\int_{-\infty}^{\infty}\left[S_{xx}(f)-\sum_{\gamma} G_{\gamma}(f) S_{xy}^{\gamma}(f-\gamma/2)^{*}\right] df.$$

In the special case for which $y(t)$ is purely stationary this general formula reduces to

$$\int_{-\infty}^{\infty} S_{xx}(f)\left[1-\sum_{\alpha}|\rho^{\alpha}(f-\alpha/2)|^{2}\right] df,$$

where $\rho^{\alpha}(f)$ is the mutual spectral coherency function

$$\rho^{\alpha}(f) \triangleq \frac{S_{xy}^{\alpha}(f)}{[S_{xx}(f+\alpha/2) S_{yy}(f-\alpha/2)]^{1/2}}.$$

It can be seen from this special case that for each value of α for which the spectral coherence is nonzero, the minimized estimation error is reduced by an amount determined by the degree of spectral coherence $|\rho^{\alpha}(f)|$.

This situation where the time-series $x(t)$ being estimated exhibits cyclostationarity but the time-series $y(t)$ used to obtain the estimate is purely stationary arises in the problem of identification of time-variant systems. If $y(t)$ is the input to an unknown system and $x(t)$ is the output, and if the time-variations of the system are periodic (or sums of periodic) functions, then we have this special case. When all of the values of α present in the system's time variations are included in the time-variant linear transformation (the system model) that produces the optimum estimate of the output $x(t)$ of the unknown system using the known input $y(t)$, then this optimum transformation is identical to the unknown system [12, Section 14C].

As an example of a signal estimation application, if $x(t)$ is the amplitude-modulated sine wave (47) and $y(t)$ is the sum of $x(t)$ and white noise, the error (65) resulting from the cyclic Wiener filter can be as small as one half its minimum obtainable without frequency shifts, i.e., using the (noncyclic) Wiener filter [21]. In this case only two shifts, $\gamma=\pm 2f_0$, are needed.

As another such example, if $x(t)$ is the amplitude-modulated pulse-train (49) in additive white noise, then for a pulse duration of T/K ($K>1$), the cyclic Wiener filtering error (65) can be as small as $1/K$ times its smallest value, the (noncyclic) Wiener filtering error, obtainable without frequency shifts [21]. More dramatic improvements in performance through the use of spectral coherence can be obtained when signal distortion is severe and also when the additive noise as well as the signal exhibits cyclostationarity. This latter case occurs, for instance, when the noise contains interfering modulated signals. For example, K amplitude-modulated pulse trains that overlap in both time and frequency and have positive-frequency bandwidth equal to $K/2$ times the pulse rate can be perfectly separated from each other [19]. Similarly, K temporally and spectrally overlapping phase-shift keyed signals with positive-frequency bandwidth equal to K times the keying rate can be perfectly separated [19], and two overlapping amplitude modulated signals can be separated, regardless of bandwidth [19].

The value of using frequency shifting in addition to frequency weighting and phase-shifting (which result from time-shifting) can be easily seen for the simple example in which interference in some portions of the signal frequency band is so strong that it overpowers the signal in those partial bands. In this case, a time-invariant filter can only reject both the signal and the interference in those highly corrupted bands, whereas a frequency-shift filter can replace the rejected spectral components with spectral components from other uncorrupted (or at least less corrupted) bands that are highly correlated with the rejected components from the signal.

Another example involves reduction of signal distortion due to frequency-selective fading caused

by multipath propagation. Straightforward amplification in faded portions of the spectrum using a time-invariant filter suffers from the resultant amplification of noise. In contrast to this, a periodically time-variant filter can replace the faded spectral components with stronger highly correlated components from other bands. If these correlated spectral components are weaker than the original components before fading, there will be some noise enhancement when they are amplified. But the amount of noise enhancement can be much less than that which would result from the time-invariant filter, which can only amplify the very weak faded components.

The relationship between frequency-shift prediction for scalar-valued purely cyclostationary time-series and conventional prediction for vector-valued purely stationary time-series is pursued in [33, 35, 39].

4.4. *Antenna array processing: spatial coherence*

Spatial coherence plays an important role when linear combining is used with an array of spatially separated single-antenna receivers in the technique of spatially diverse communications which is used to combat signal fading due to propagation phenomena [46]. Also, spatial coherence is the main property that is used by memoryless antenna arrays to form spatial reception beams and nulls (spatial filters) in order to extract a desired signal while rejecting interfering signals impinging on the array [10]. The array can be viewed as providing spatial samples of a space–time waveform, and the spatial coherence is used by forming linear combinations of the spatial samples. Also, algorithms for adaptively adjusting the weights in the linear combination typically exploit the spatial coherence. This is true for signals that adjust the array to extract signals as well as algorithms that just measure signal parameters such as direction of arrival (DOA). The currently popular (among investigators in the algorithms research community) spatially-coherent-signal subspace algorithms MUSIC and ESPRIT are examples of this [36, 43, 45].

To briefly illustrate the idea, we consider an array of n antenna elements together with n corresponding receivers that produce an n-vector $\boldsymbol{y}(t) = [y(t, \zeta_1), y(t, \zeta_2), \ldots, y(t, \zeta_n)]$ of analytic signals, where $\{\zeta_1, \zeta_2, \ldots, \zeta_n\}$ represents the spatial locations of the n sensors. These signals can often be modeled using a narrowband approximation as follows:

$$\boldsymbol{y}(t) = \boldsymbol{A}\boldsymbol{s}(t) + \boldsymbol{n}(t), \tag{67}$$

where $\boldsymbol{A}$ is an $n \times d$ matrix whose columns $\{\boldsymbol{a}(\theta_i): i = 1, 2, \ldots, d\}$ are the direction vectors associated with the d signals [the elements of $\boldsymbol{s}(t)$] impinging on the array; $\{\theta_i: i = 1, 2, \ldots, d\}$ are the corresponding d DOAs; and $\boldsymbol{n}(t)$ is additive zero-mean sensor noise, assumed to be independent from sensor to sensor, and of equal average power σ^2 for all sensors. It follows from this model that the $n \times n$ matrix of temporal crosscorrelations of the elements of $\boldsymbol{y}(t)$ is given by

$$\boldsymbol{R}_{yy} \triangleq \langle \boldsymbol{y}(t)\boldsymbol{y}^\dagger(t) \rangle = \boldsymbol{A}\boldsymbol{R}_{ss}\boldsymbol{A}^\dagger + \sigma^2\boldsymbol{I},$$

where $\boldsymbol{A}^\dagger$ is the transpose conjugate of $\boldsymbol{A}$ and $\boldsymbol{I}$ is the $n \times n$ identity matrix. The spatial coherence of the d signals is reflected in the structure of this spatial correlation matrix and this structure can be used for purposes of estimating DOAs and extracting signals.

For example, the rank of the $n \times n$ matrix $\boldsymbol{A}\boldsymbol{R}_{ss}\boldsymbol{A}^\dagger$ is no more than the rank of the $d \times d$ matrix $\boldsymbol{R}_{ss}$, which cannot exceed its dimension d. If $d < n$ and no one of the d signals in $\boldsymbol{s}(t)$ is completely correlated with any of the other $d-1$ signals, so that $\boldsymbol{R}_{ss}$ is of full rank, then the null space of $\boldsymbol{A}\boldsymbol{R}_{ss}\boldsymbol{A}^\dagger$, which is orthogonal to all d direction vectors $\{\boldsymbol{a}(\theta_i)\}$, is spanned by the eigenvectors corresponding to the $n-d$ smallest eigenvalues of $\boldsymbol{R}_{yy}$ (each of which equals σ^2). Thus, singular value decomposition of $\boldsymbol{R}_{yy}$ can be used to estimate the number d of signals, and to estimate the d DOAs by searching over θ for the vectors $\boldsymbol{a}(\theta)$ (assuming the array calibration function $\boldsymbol{a}(\cdot)$ is known) that are most nearly orthogonal to all of the eigenvectors corresponding to the $n-d$ smallest eigenvalues.

For broadband signals, temporal coherence, as well as spatial coherence, can be exploited by using arrays with memory (cf. [49]).

4.5. Antenna array processing: spectral and spatial coherence

When the signals impinging on an antenna array are modulated sine waves and/or periodic pulse-trains, their spectral coherence, as well as their spatial coherence (and also temporal coherence for broadband signals), can be exploited for signal parameter estimation and signal extraction.

To illustrate how spectral coherence can be used, we consider the situation in which $d_\alpha \leq d$ of the d signals impinging on the array all exhibit cyclostationarity with the same cycle frequency α (assumed known). In this case, the cyclic spatial correlation matrix

$$\langle \boldsymbol{R}^{\alpha}_{yy}(\tau)\rangle \triangleq \langle \boldsymbol{y}(t+\tau/2)\boldsymbol{y}^{\dagger}(t-\tau/2)\, e^{-i2\pi\alpha t}\rangle$$

for the model (67) takes the form

$$\boldsymbol{R}^{\alpha}_{yy}(\tau) = \boldsymbol{A}\boldsymbol{R}^{\alpha}_{ss}(\tau)\boldsymbol{A}^{\dagger},$$

where $\boldsymbol{A}$ is the matrix of direction vectors for the d_α signals and $\boldsymbol{R}^{\alpha}_{ss}(\tau)$ is their $d_\alpha \times d_\alpha$ correlation matrix. Since there is no contribution to $\boldsymbol{R}^{\alpha}_{yy}(\tau)$ from the sensor noise or the $d-d_\alpha$ signals that do not exhibit cyclostationarity with cycle frequency α (or spectral coherence with frequency separation α), we have performed a signal selection or sorting function by simply obtaining a reliable measurement of this matrix. This matrix is independent of the particular characteristics of the sensor noise and the number and characteristics of interfering signals. If $\boldsymbol{R}^{\alpha}_{ss}(\tau)$ is of full rank, then the d_α direction vectors $\{\boldsymbol{a}(\theta_i)\}$ are all orthogonal to the null space of $\boldsymbol{R}^{\alpha}_{yy}(\tau)$, assuming $d_\alpha < n$. Therefore, the d_α DOAs $\{\theta_i\}$ can be estimated by searching over θ for the vectors $\alpha(\theta)$ that are most nearly orthogonal to all basis vectors in this null space. To find the approximate null space, we can use the usual singular value decomposition methods. The effectiveness of such spectral coherence exploiting methods of DOA estimation is demonstrated in [15, Section 12.8; 41, 44]. For example, if $d_\alpha = 1$ signal having cycle frequency α arrives from an angle θ that is very close to the angle from which an interfering signal arrives, then the spectral-coherence-exploiting methods of DOA estimation need only estimate the angle θ of the desired signal, whereas a conventional method such as MUSIC or ESPRIT must successfully resolve the angles of both signals, which typically requires much greater data collection time.

In addition, the concept of Spectral COherence REstoral (SCORE) can be used to develop algorithms that blindly (without the use of training signals, knowledge of DOA, knowledge of array geometry or array calibration) adapt antenna arrays to extract desired signals while rejecting undesired interfering signals (i.e., to spatially filter the received data) [1, 2; 15, Section 12.8; 40]. Spectral coherence restoral exploits the spectral coherence property of a signal of interest by using an appropriate frequency shift α that preserves correlation for the signal of interest while decorrelating interfering signals and noise. This is analogous to the adaptive spectral-line-enhancement filtering technique that uses a time shift to preserve temporal correlation for narrowband signal components while decorrelating wideband noise components.

As an example of a SCORE technique, the vector $\boldsymbol{w}$ of weights in the linear combination $\boldsymbol{w}^{\dagger}\boldsymbol{y}(t)$ of the spatial samples the effectively forms an antenna reception pattern can be taken to be that vector which jointly with the vector $\boldsymbol{v}$ maximizes the degree of mutual spectral coherence

$$\rho = \frac{R^{\alpha}_{xz}(\tau)}{[R_{xx}(0)R_{zz}(0)]^{1/2}} = \frac{\boldsymbol{w}^{\dagger}\boldsymbol{R}^{\alpha}_{yy}(\tau)\boldsymbol{v}}{[(\boldsymbol{w}^{\dagger}\boldsymbol{R}_{yy}\boldsymbol{w})(\boldsymbol{v}^{\dagger}\boldsymbol{R}_{yy}\boldsymbol{v})]^{1/2}}$$

between the linearly combined signals

$$x(t) = \boldsymbol{w}^{\dagger}\boldsymbol{y}(t) \quad \text{and} \quad z(t) = \boldsymbol{v}^{\dagger}\boldsymbol{y}(t).$$

When there are d_α signals in $\boldsymbol{y}(t)$ exhibiting cyclostationarity with cycle frequency α, then the first signal is still estimated by maximizing the mutual spectral coherence ρ. The second signal is estimated by maximizing the partial coherence (e.g. see Subsection 2.3) between $x_2(t)$ and $z_2(t)$ after the

effects of $x_1(t)$ and $z_1(t)$ have been removed from $y(t)$ and $y(t-\tau)\,e^{i2\pi\alpha t}$, respectively. Similarly, the lth signal is estimated by maximizing the partial coherence between $x_l(t)$ and $z_l(t)$ after the effects of $x_1(t), \ldots, x_{l-1}(t)$ and $z_1(t), \ldots, z_{l-1}(t)$ have been removed from $y(t)$ and $y(t-\tau)\,e^{i2\pi\alpha t}$, respectively. The d_α solutions $\boldsymbol{w}$ are given by the generalized eigenvectors corresponding to the d_α nonzero generalized eigenvalues of the generalized eigenequation [15, Section 12.8]

$$\boldsymbol{R}^{\alpha}_{yy}(\tau)\boldsymbol{R}^{-1}_{yy}\boldsymbol{R}^{\alpha}_{yy}(\tau)^{\dagger}\boldsymbol{w}=\lambda\boldsymbol{R}_{yy}\boldsymbol{w}. \tag{68}$$

Each solution vector $\boldsymbol{w}$ is optimal for one of the d_α signals. This SCORE technique can be derived as a particular application of common factor analysis (also called canonical correlation analysis), which is well known in the multivariate statistics community [30]. Since only the d_α signals in $\boldsymbol{y}(t)$ exhibiting cyclostationarity with cycle frequency α are correlated with components in $\boldsymbol{y}(t-\tau)\,e^{i2\pi\alpha t}$, the adaptive spatial filtering problem can be re-interpreted as the common factor analysis problem in which it is desired to estimate the signal components that are common to two data sets. By choosing these two data sets to be $\boldsymbol{y}(t)$ and $\boldsymbol{y}(t-\tau)\,e^{i2\pi\alpha t}$ and applying standard results from common factor analysis, the algorithm (68) is obtained.

Another SCORE technique uses the generalized eigenvectors of the simpler generalized eigenequation [40]

$$\boldsymbol{R}^{\alpha}_{yy}(\tau)\boldsymbol{w}=\lambda\boldsymbol{R}_{yy}\boldsymbol{w}.$$

Blind adaptive SCORE techniques, which exploit both spectral (through $\boldsymbol{R}^{\alpha}_{yy}$) and spatial (through $\boldsymbol{R}^{\alpha}_{yy}$ and $\boldsymbol{R}_{yy}$) coherence, often produce nearly (or exactly) maximum signal-to-interference-and-noise-ratio solutions [2]. For example, if there is only $d_\alpha = 1$ signal of interest impinging on the array that exhibits cyclostationarity with cycle frequency α, then the solution corresponding to the one nonzero eigenvalue of (68) is given by

$$\boldsymbol{w}\propto\boldsymbol{R}^{-1}_{yy}\boldsymbol{a}(\theta_1),$$

which is the well-known maximum-SINR (and minimum-mean-squared-error) weight vector for a signal arriving from direction θ_1 for any array and any number and type of interfering signals and noises that are uncorrelated with the signal of interest.

4.6. *Detection and classification: spectral and temporal coherence*

Randomly fluctuating modulated signals that are severely masked by other interfering signals and noise can be more effectively detected, in some applications, by detection of spectral correlation rather than detection of energy. This is so, for instance, when the energy level of the background noise or interference fluctuates unpredictably, thereby complicating the problem of setting energy threshold levels [12, Section 14E; 14; 15, Section 12.8; 22]. The utility of spectral correlation for detection is illustrated in Fig. 4, where it is difficult to see the contribution from the signal in the measured spectrum at $\alpha=0$, but it is easy to see the contribution in the measured spectral correlation function at $\alpha=\pm 2f_0$, $\alpha=\pm 1/T$, $\alpha=\pm 2f_0 \pm 1/T$. The signal here is a binary phase-shift-keyed sine wave (BPSK)[11] with sine-wave frequency f_0 and

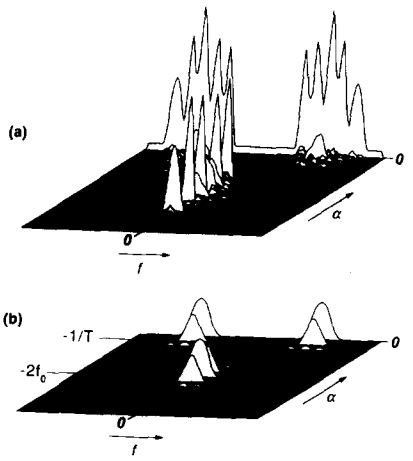

Fig. 4. Measured spectral correlation magnitudes for simulated signals: (a) BPSK signal corrupted by noise and interference; (b) uncorrupted BPSK signal.

[11] The BPSK signal can be obtained by forming the product of a sine wave with a binary valued pulse-amplitude modulated signal of the form (49).

keying interval T, the noise is white, and the interference consists of five amplitude-modulated sine waves. Figure 4(a) shows the measured spectral correlation magnitude surface for the corrupted BPSK signal, and Fig. 4(b) shows the measured surface for the uncorrupted BPSK signal.

This example also suggests that the unique patterns of spectral correlation that are exhibited by different types of modulated signals (cf. [12, Chapter 12; 15, Chapter 12]) can be effectively exploited for recognition of modulation type for severely corrupted signals. However, the production of reliable spectral correlation surfaces is a computationally intensive task. A survey of methods of measurement of spectral correlation is given in [12, Chapter 13] and computationally efficient digital algorithms are presented in [6, 38].

To illustrate the potential for using spectral correlation characteristics to classify signals, the spectral correlation surfaces for three types of PSK signals with keying rate $1/T$ are shown in Fig. 5. The signal types are BPSK, quaternary PSK (QPSK) and staggered QPSK (SQPSK). It can be seen that the spectra ($\alpha = 0$) of these three signals are identical to each other and therefore cannot be used to distinguish among these signal types. However, the overall spectral correlation surfaces are highly distinct and can therefore be used effectively for classification.

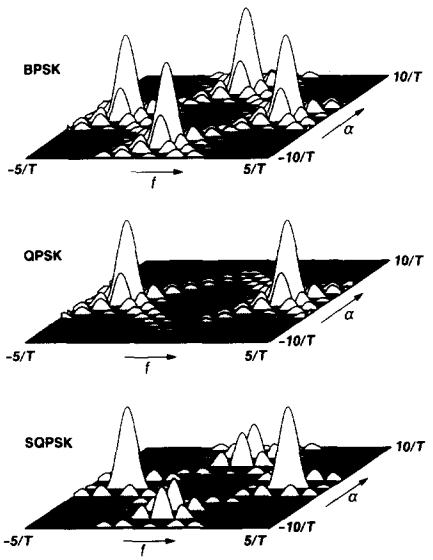

Fig. 5. Theoretical spectral correlation magnitudes for three types of PSK signals: BPSK, QPSK and SQPSK.

4.7. TDOA estimation: spectral and temporal coherence

Parameters such as time-difference-of-arrival (TDOA) of signals at two sensors, which can be used for radiating-signal source location, can be more effectively estimated for modulated signals in some applications by using measurements of cyclic ($\alpha \neq 0$) cross-correlations and cyclic cross-spectra, rather than conventional ($\alpha = 0$) cross-correlations and cross-spectra. This is so, for example, when there is a multiplicity of interfering modulated signals that cannot be separated because of the lack of a sufficient number of sensors, such as antenna elements, or when the signal is weak relative to the noise or interference [9, 20]. This can be seen as follows. The two signals received at two sensors can be modeled as

$$y_1(t) = s(t) + n(t),$$
$$y_2(t) = (a)s(t - t_0) + m(t), \tag{69}$$

where t_0 is the TDOA of the signal $s(t)$, and $n(t)$ and $m(t)$ consist of interfering zero-mean signals and noise that are independent of $s(t)$. For this model the ratio of the cross-spectrum of $y_1(t)$ and $y_2(t)$ to the auto-spectrum of $y_1(t)$ is given by

$$\frac{S_{21}(f)}{S_{11}(f)} = \frac{(a)\exp(-i2\pi f t_0)S_{ss}(f) + S_{mn}(f)}{S_{ss}(f) + S_{nn}(f)}.$$

In the absence of noise and especially interference, $S_{nm}(f) \equiv S_{nn}(f) \equiv 0$ and therefore

$$\frac{S_{21}(f)}{S_{11}(f)} = (a)\exp(-i2\pi f t_0), \quad f \in B,$$

where B is the support of $S_{ss}(f)$. Thus, the inverse Fourier transform of this ratio of spectra peaks at t_0 (cf. [8]). However, when $S_{nm}(f)$ is not negligible,

the desired linear (in f) phase characteristic of this ratio of spectra can be severely masked. This is illustrated with the auto-spectral correlation magnitude $|S_{11}^{\alpha}(f)|$ shown in Fig. 4 (for the same signal, interference and noise environment as that considered in the preceding subsection on detection and classification), which closely resembles the cross-spectral correlation magnitude $|S_{12}^{\alpha}(f)|$. (Since the magnitude of the signal component of the cross-spectrum at $\alpha = 0$ is completely masked by the interference components, its phase characteristic is also severely corrupted.) Notice, however, that the cross spectral correlation at any of the cycle frequencies α of the BPSK is relatively free of contributions from the noise and interference. In fact, for the model (69), we have

$$\frac{S_{21}^{\alpha}(f)}{S_{11}^{\alpha}(f)} = \frac{(a)\exp(-\mathrm{i}[2\pi f t_0 + \theta])S_{ss}^{\alpha}(f) + S_{mn}^{\alpha}(f)}{S_{ss}^{\alpha}(f) + S_{nn}^{\alpha}(f)},$$

where $\theta = \pi\alpha\tau_0$. Thus, as long as the noise and interference do not exhibit spectral coherence with frequency separation α corresponding to a cycle frequency of the signal $s(t)$, then $S_{mn}^{\alpha}(f) \equiv S_{nn}^{\alpha}(f) \equiv 0$ and $S_{ss}^{\alpha}(f) \neq 0$ for (say) $f \in B_{\alpha}$ and, consequently,

$$\frac{S_{21}^{\alpha}(f)}{S_{11}^{\alpha}(f)} = (a)\exp(-\mathrm{i}[2\pi f t_0 + \theta]), \quad f \in B_{\alpha},$$

as desired. The effectiveness of such spectral-coherence exploiting TDOA estimation methods is demonstrated in [9; 15, Section 12.8; 20].

5. Summary

In this paper, a unifying view of the concept of coherence is presented. Precise mathematical definitions of the degree of coherence of various types, including temporal, spectral and spatial coherence as well as self, mutual and partial coherence, are given. A spectral decomposition of coherence is developed, and the connection between cyclostationarity and spectral coherence is explained. The practical limitations of the somewhat evasive property of spectral coherence of nonstationary/non-cyclostationary processes are revealed. Geometric interpretations of coherence and its intimate link with linear relationships are discussed. A variety of physical meanings of the term coherence and a broad sampling of uses of coherence properties in signal processing are briefly described.

Of the three types of coherence, temporal, spatial and spectral, the third type has not received as much attention in the literature as the first two. This is due in part to the fact that only signals that exhibit cyclostationarity or local cyclostationarity (or nonstationarity of a form that is known prior to measurement) possess useful properties of spectral coherence (when ensemble averaging is not possible). To help promote awareness of the utility of spectral coherence, a little more emphasis is placed on the uses of this particular type of coherence. A more detailed tutorial treatment of spectral coherence is given in [16].

It is hoped that the unifying view of the fundamental concept of coherence presented in this paper will enhance understanding and foster uses of coherence throughout the broad field of signal processing. Although the treatment of coherence presented here is intended to be uniquely comprehensive and unifying, the scope is still limited. Topics omitted, including primarily relationships between coherence and causality (cf. [7, 25, 37]), methods of and performance evaluation for coherence measurement (cf. [6, 12]), physical modeling of coherence in the study of wave propagation and scattering, and various physical meanings and uses of coherence, are not necessarily less important or in any other way subordinate to those included.

References

[1] B.G. Agee, S.V. Schell and W.A. Gardner, "The SCORE approach to blind adaptive signal extraction: An application of the theory of spectral correlation", *Proc. IEEE/ASSP Fourth Workshop on Spectrum Estimation and Modeling*, Minneapolis, MN, 3-5 August 1988, pp. 277–282.

[2] B.G. Agee, S.V. Schell and W.A. Gardner, "The self-coherence restoral (SCORE) approach to blind adaptive array processing", *Proc. IEEE (Special Issue on Multidimensional Signal Processing)*, Vol. 78, No. 4, April 1990, pp. 753–767.

[3] J. Bass, "Stationary functions and their applications to the theory of turbulence. I. Stationary functions; II. Turbulent solutions of the Navier Stokes Equations", *J. Math. Anal. Appl.*, Vol. 47, 1974, pp. 354–399 and 458–503.

[4] M. Born and E. Wolf, *Principles of Optics*, Pergamon, New York; 1959.

[5] R.N. Bracewell, "Defining the coherence of a signal", *Proc. IRE*, Vol. 50, February 1962, p. 214.

[6] W.A. Brown and H.H. Loomis, Jr., "Digital implementations of spectral correlation analyzers", *IEEE Trans. Signal Process.*, Vol. 40, 1992 (in press).

[7] J.A. Cadzow and O.M. Solomon, Jr., "Linear modeling and the coherence function", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-35, January 1987, pp. 19–28.

[8] G.C. Carter, "Coherence and time delay estimation", *Proc. IEEE*, Vol. 75, No. 2, 1987, pp. 236–255.

[9] C.K. Chen and W.A. Gardner, "Signal-selective time-difference-of-arrival estimation for passive location of manmade signal sources in highly corruptive environments. Part II: Algorithms and performance", *IEEE Trans. Signal Process.*, Vol. 40, 1992, pp. 1185–1197.

[10] R.T. Compton, Jr., *Adaptive Antennas*, Prentice Hall, Englewood Cliffs, NJ, 1988.

[11] I. Daubechies, "The wavelet transform, time-frequency localization and signal-analysis', *IEEE Trans. Inform. Theory*, Vol. 36, No. 5, September 1990, pp. 961–1005.

[12] W.A. Gardner, *Statistical Spectral Analysis: A Nonprobabilistic Theory*, Prentice Hall, Englewood Cliffs, NJ, 1987.

[13] W.A. Gardner, "Correlation estimation and time-series modeling for nonstationary processes", *Signal Processing*, Vol. 15, No. 1, July 1988, pp. 31–41.

[14] W.A. Gardner, "Signal interception: A unifying theoretical framework for feature detection", *IEEE Trans. Comm.*, Vol. COM-36, No. 8, 1988, pp. 897–906.

[15] W.A. Gardner, *Introduction to Random Processes with Applications to Signals and Systems*, Macmillan, New York, 1985, Second Edition, McGraw-Hill, New York, 1989.

[16] W.A. Gardner, "Exploitation of spectral redundancy in cyclostationary signals", *IEEE Signal Process. Mag.*, Vol. 8, April 1991, pp. 14–36.

[17] W.A. Gardner, "Two alternative philosophies for estimation of the parameters of time-series", *IEEE Trans. Inform. Theory*, Vol. 37, 1991, pp. 216–218.

[18] W.A. Gardner, "On the spectral coherence of nonstationary processes", *IEEE Trans. Signal Process.*, Vol. 39, 1991, pp. 424–430.

[19] W.A. Gardner, "Cyclic Wiener filtering: Theory and method", *IEEE Trans. Comm.*, 1992 (in press).

[20] W.A. Gardner and C.K. Chen, "Signal-selective time-difference-of-arrival estimation for passive location of manmade signal sources in highly corruptive environments. Part I: Theory and method", *IEEE Trans. Signal Process*, Vol. 40, 1992, pp. 1168–1184.

[21] W.A. Gardner and L.E. Franks, "Characterization of cyclostationary random signal processes", *IEEE Trans. Inform. Theory*, Vol. IT-21, 1975, pp. 4–14.

[22] W.A. Gardner and C.M. Spooner, "Signal interception: Performance advantages of cyclic feature detectors". *IEEE Trans. Comm.*, Vol. 40, 1992, pp. 149–159.

[23] R.J. Glauber, "The quantum theory of optical coherence", *Phys. Rev.*, Vol. 130, No. 6, June 1963, pp. 2529–2539.

[24] R.J. Glauber, "Coherent and incoherent states of the radiation field", *Phys. Rev.*, Vol. 131, No. 6, September 1963, pp. 2766–2788.

[25] C.W.J. Granger, "Investigating causal relations by econometric models and cross-spectral methods", *Econometrica*, Vol. 37, No. 3, 1969, pp. 424–438.

[26] K. Hasselmann and T.P. Barnett, "Techniques for linear prediction for systems with periodic statistics", *J. Atmospheric Sci.*, Vol. 38, 1981, pp. 2275–2283.

[27] C.W. Helstrom, "An expansion of a signal in Gaussian elementary signals", *IEEE Trans. Inform. Theory*, Vol. 12, January 1966, pp. 81–82.

[28] A. Ishimaru, *Wave Propagation and Scattering in Random Media*, Academic Press, New York, 1978.

[29] G.M. Jenkins and D.G. Watts, *Spectral Analysis and its Applications*, Holden-Day, San Francisco, CA, 1968.

[30] R.A. Johnson and D.W. Wichern, *Applied Multivariate Statistical Analysis*, Second Edition, Prentice Hall, Englewood Cliffs, NJ, 1988.

[31] L. Mandel and E. Wolf, "Coherence properties of optical fields", *Rev. Modern Phys.*, Vol. 37, No. 2, April 1965, pp. 231–287.

[32] L. Mandel and E. Wolf, "Spectral coherence and the concept of cross-spectral purity", *J. Opt. Soc. Amer.*, Vol. 66, June 1976, pp. 529–535.

[33] A.G. Miamee and H. Salehi, "On the prediction of periodically correlated stochastic processes," in: P.R. Krishnaiah, ed., *Multivariate Analysis V*, North-Holland, Amsterdam, 1980, pp. 167–179.

[34] W.H. Munk, F.E. Snodgrass and M.J. Tucker, "Spectra of low-frequency ocean waves", *Bull. Scripps Inst. Oceanography*, Vol. 7, 1959, pp. 283–362.

[35] M. Pagano, "On periodic and multiple autoregressions", *Ann. Statist.*, Vol. 6, No. 6, 1978, pp. 1310–1317.

[36] A. Paulraj, R. Roy and T. Kailath, "Estimation of signal parameters via rotational invariance techniques – ESPRIT", *Conf. Record, Nineteenth Asilomar Conf. on Circuits, Systems, and Computers*, Pacific Grove, CA, 6–8 November 1985, pp. 83–89.

[37] D.A. Pierce and L.D. Haugh, "Causality in temporal systems", *J. Econometrics*, Vol. 5, 1977, pp. 265–293.

[38] R.S. Roberts, W.A. Brown and H.H. Loomis, "Computationally efficient algorithms for cyclic spectral analysis", *IEEE Signal Process. Mag.*, Vol. 8, April 1991, pp. 38–49.

[39] H. Sakai, "Circular lattice filtering using Pagono's method", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-30, No. 2, April 1982, pp. 279–287.

[40] S.V. Schell and B.G. Agee, "Application of the SCORE algorithm and SCORE extensions to sorting in the rank-L spectral self-coherence environment", *Proc. Twenty-Second Annual Asilomar Conf. on Signals, Systems, and Computers*, Pacific Grove, CA, 30 October–2 November 1988.

[41] S.V. Schell and W.A. Gardner, "Signal-selective high-resolution direction finding in multipath", *Proc. Internat. Conf. Acoust. Speech Signal Process. 1990*, Albuquerque, NM, 3–6 April 1990, pp. 2667–2670.

[42] S.V. Schell and W.A. Gardner, "Progress on signal-selective direction finding", *Proc. Fifth IEEE/ASSP Workshop on Spectrum Estimation and Modeling*, Rochester, NY, 10–12 October 1990, pp. 144–148.

[43] S.V. Schell and W.A. Gardner, "High-resolution direction finding", Chapter in: N.K. Bose and C.R. Rao, eds., *Handbook of Statistics*, Vol. 10, Elsevier, Amsterdam, 1992 (in press).

[44] S.V. Schell, R.A. Calabretta, W.A. Gardner and B.G. Agee, "Cyclic MUSIC algorithms for signal-selective direction estimation", *Proc. Internat. Conf. Acoust. Speech Signal Process. 1989*, Glasgow, 23–26 May 1989, pp. 2278–2281.

[45] R.O. Schmidt, A signal subspace approach to multiple emitter location and spectral estimation, Ph.D. Dissertation, Department of Electrical Engineering, Stanford University, Stanford, CA, 1981.

[46] M. Schwartz, W.R. Bennett and S. Stein, *Communication Systems and Techniques*, McGraw-Hill, New York, 1966, Chapter 10.

[47] A.E. Siegman, *Lasers*, University Science Books, Mill Valley, CA, 1986.

[48] A.R. Thompson, J.M. Moran and G.W. Swenson, Jr., *Interferometry and Synthesis in Radio Astronomy*, Wiley, New York, 1986.

[49] H. Wang and M. Kaveh, "Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources", *IEEE Trans. Acoust. Speech Signal Process.*, Vol. ASSP-33, August 1985, pp. 823–831.

[50] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, The Technology Press of MIT and John Wiley & Sons, New York, 1949. Originally issued in February 1942, as a classified National Defense Council Research Report.